

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2001, 21:49:06 ; Search time 1167.42 Seconds
(without alignments)
7011.240 Million cell updates/sec

Title: US-09-382-242-23
Perfect score: 555
Sequence: 1 ATGCTTTAAACAAGCAGCTC.....AATCATACACTAAACGATAG 555

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1283235 seqs, 7373929652 residues
Total number of hits satisfying chosen parameters: 2566470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da1:*
2: gb_da2:*
3: gb_da3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_da1:*
17: em_da2:*
18: em_fun:*
19: em_hcgo_hum:*
20: em_hcgo_inv:*
21: em_hcgo_rnd:*
22: em_hcgo_hum1:*
23: em_hcgo_hum2:*
24: em_hcgo_hum3:*
25: em_hcgo_hum4:*
26: em_hcgo_hum5:*
27: em_hcgo_hum6:*
28: em_hcgo_hum7:*
29: em_hcgo_hum8:*
30: em_hcgo_inv1:*
31: em_hcgo_inv2:*
32: em_hcgo_inv3:*
33: em_hcgo_inv4:*
34: em_hcgo_inv5:*
35: em_hcgo_inv6:*
36: em_hcgo_inv7:*
37: em_hcgo_inv8:*
38: em_hcgo_inv9:*
39: em_hcgo_inv10:*
40: em_hcgo_inv11:*
41: em_hcgo_inv12:*
42: em_hcgo_inv13:*
43: em_hcgo_inv14:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_v1:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_v11:*
59: gb_v12:*
60: gb_v13:*
61: gb_v14:*
62: gb_v15:*
63: gb_v16:*
64: gb_v17:*
65: gb_v18:*
66: gb_v19:*
67: gb_v20:*
68: gb_v21:*
69: gb_v22:*
70: gb_v23:*
71: gb_v24:*
72: gb_v25:*
73: gb_v26:*
74: gb_v27:*
75: gb_v28:*
76: gb_v29:*
77: gb_v30:*
78: gb_v31:*
79: gb_v32:*
80: gb_v33:*
81: gb_v34:*
82: gb_v35:*
83: gb_v36:*
84: gb_v37:*
85: gb_v38:*
86: gb_v39:*
87: gb_v40:*
88: gb_v41:*
89: gb_v42:*
90: gb_v43:*
91: gb_v44:*
92: gb_v45:*
93: gb_v46:*
94: gb_v47:*
95: gb_v48:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	48.2	8.7	217412	69	AC0244400
C 2	47.8	8.6	14001	6	PF00001R
C 3	47.6	8.6	192929	60	AC005505
C 4	47.6	8.6	256172	60	AC005139
C 5	47.6	8.6	310779	60	AC005140
C 6	47.2	8.5	108908	6	PF00001R
C 7	47.2	8.5	298469	4	AE003846
C 8	45.6	8.2	21595	64	AC014617
C 9	45.6	8.2	193742	82	AL513530
C 10	45.2	8.1	3171	6	DD068754
C 11	45	8.1	179401	62	AC010903

C 12	44.8	8.1	84346	67	AC021920	AC021920 Homo sapi
C 13	44.4	8.0	27694	71	HASMT	D31785 Pichia cana
C 14	44.4	8.0	185926	14	AC027632	AC027632 Homo sapi
C 15	44.2	8.0	118313	85	AC003991	AC003991 Human BAC
C 16	44.2	8.0	204652	84	PFMA13P6	AL049183 Plasmid
C 17	44.4	7.9	1084	53	CNS071NH	AL425139 clone BAO
C 18	44.4	7.9	45565	8	FR151J19	AL021531 Fugu rubr
C 19	44.4	7.9	138094	70	AC026839	AC026839 Homo sapi
C 20	43.8	7.9	13859	4	AE001366	AC001366 Plasmodu
C 21	43.8	7.9	84346	67	PECOMP1R	AC021920 Homo sapi
C 22	43.6	7.9	14001	6	PECOMP1R	X95276 P. falciparu
C 23	43.6	7.9	15421	6	PECOMP1R	X95275 P. falciparu
C 24	43.4	7.8	1847	5	AF209925	AF209925 Plasmodu
C 25	43.4	7.8	2787	6	PFSTARE	Z26314 P. falciparu
C 26	43.2	7.8	5708	13	AF271619	AF271619 Aspergill
C 27	43.2	7.8	16932	12	AF083031	AF083031 Guillard
C 28	43.2	7.8	169546	60	AC004157	AC004157 Plasmodu
C 29	43	7.7	34996	6	U87145	U87145 T.oxoplasma
C 30	43	7.7	171798	4	AC084397	AC084397 Trypanoso
C 31	42.8	7.7	840	14	CNS0180K	AL110675 Bolyalis
C 32	42.8	7.7	149789	60	AC007926	AC007926 Trypanoso
C 33	42.8	7.7	149789	60	AC007926	AC007926 Trypanoso
C 34	42.8	7.7	180290	85	AC005081	AC005081 Homo sapi
C 35	42.6	7.7	1074	5	AF190049	AF190049 Pterostic
C 36	42.6	7.7	94384	87	AC011718	AC011718 Homo sapi
C 37	42.6	7.7	104992	60	AC005504	AC005504 Plasmodu
C 38	42.6	7.7	177951	76	AC074383	AC074383 Homo sapi
C 39	42.6	7.7	181980	63	AC013394	AC013394 Homo sapi
C 40	42.4	7.6	270	13	AY021751	AY021751 Oryza sat
C 41	42.2	7.6	801	53	CNS06EMJ	AL395297 T7 end of
C 42	42.2	7.6	1051	53	CNS06EMJ	AL392832 T3 end of
C 43	42.2	7.6	5769	5	AF104350	AF104350 Dictyoste
C 44	42.2	7.6	8212	43	ALR1BPR	X75653 Astasia lon
C 45	42.2	7.6	8307	4	AC024860	AC024860 Caenorhab

ALIGNMENTS

RESULT 1	AC024400/c	217412 bp	DNA	HTG	13-JUL-2000
LOCUS	Homo sapiens clone Rp11-439d				
DEFINITION	AC024400.2	GI:7767819			
ACCESSION	AC024400.2	GI:7767819			
VERSION	AC024400.2	GI:7767819			
KEYWORDS	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 217412)				
AUTHORS	Bliren, B., Linton, L., Nusbaum, C., Lander, E., Abrahams, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bede, F., Boguski, L., Bouck, B., Brown, A., Burkett, G., Campilano, A., Castle, A., Chong, Y., Collinge, M., Collins, S., Collins, A., Cooke, P., Dear, R., Dewar, K., Diaz, J., Dodge, S., Domingo, M., Doyle, M., Fagan, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J., Gardy, S., Ginde, S., Goyette, M., Graham, L., Grand, P., Grant, G., Hago, B., Heath, A., Horton, L., Howland, J., Iliev, I., Johnson, R., Jones, C., Kahn, L., Karitas, A., Klein, J., Landers, T., Laroque, K., Lech, J., Levine, R., Meneis, L., Mihov, T., Miranda, C., Mlepe, V., Morrow, J., Naylor, J., Norman, C., O'Connor, T., O'Donnell, P., O'Neill, D., Olyar, T., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schaner, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, J., Subramanian, A., Talamas, J., Testa, S., Theodore, J., Tiller, A.,				

TITLE
JOURNAL
COMMENT

Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Submitted (28-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On May 10, 2000 this sequence version replaced g1:7108401.

All repeats were identified using RepeatMasker:

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: W18R

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L6776

Center clone name: 439_D_4

NOTE: This record contains 242 individual sequencing reads that have not been assembled into contigs. Runs of N are used to separate the reads and the order in which they appear is completely arbitrary. Low-pass sequence sampling is useful for identifying clones that may be gene-rich and allows overlap relationships among clones to be deduced. However, it should not be assumed that this clone will be sequenced to completion. In the event that the record is updated, the accession number will be preserved.

1 716 815: gap of 100 bp in length

816 1554: contig of 739 bp in length

1555 1654: gap of 100 bp

1655 2421: contig of 767 bp in length

2422 2521: gap of 100 bp

2522 3271: contig of 750 bp in length

3272 3371: gap of 100 bp

3372 4116: contig of 745 bp in length

4117 4216: gap of 100 bp

4217 4979: contig of 763 bp in length

4980 5079: gap of 100 bp

5080 5835: contig of 760 bp in length

5840 5939: gap of 100 bp

5940 6709: contig of 770 bp in length

6710 6809: gap of 100 bp

6810 7579: contig of 770 bp in length

7580 7679: gap of 100 bp

7680 8381: contig of 702 bp in length

8382 8481: gap of 100 bp

8482 9219: contig of 738 bp in length

9220 9319: gap of 100 bp

9320 10072: contig of 753 bp in length

10073 10172: gap of 100 bp

10173 10927: contig of 755 bp in length

10928 11027: gap of 100 bp

11028 11797: contig of 770 bp in length

11798 11897: gap of 100 bp

11898 12659: contig of 762 bp in length

12660 12759: gap of 100 bp

12760 13535: contig of 776 bp in length

13536 13635: gap of 100 bp

13636 14355: contig of 720 bp in length

14356 14455: gap of 100 bp

14456 15218: contig of 763 bp in length

15219 15318: gap of 100 bp

15319 16081: contig of 763 bp in length

16082 16181: gap of 100 bp

16182 16943: contig of 762 bp in length

16944 17043: gap of 100 bp

17044 17168: contig of 725 bp in length

17169 17668: gap of 100 bp

17669 18613: contig of 745 bp in length

[illegible]

*	49468	50197:	contig of 730 bp	in length
*	50198	50297:	gap of 100 bp	
*	50298	51049:	contig of 752 bp	in length
*	51050	51149:	gap of 100 bp	
*	51150	51907:	contig of 758 bp	in length
*	51908	52007:	gap of 100 bp	
*	52008	52728:	contig of 721 bp	in length
*	52729	52828:	gap of 100 bp	
*	52829	53594:	contig of 766 bp	in length
*	53595	53694:	gap of 100 bp	
*	53695	54432:	contig of 738 bp	in length
*	54433	54532:	gap of 100 bp	
*	54533	55295:	contig of 763 bp	in length
*	55296	55395:	gap of 100 bp	
*	55396	56159:	contig of 764 bp	in length
*	56160	56259:	gap of 100 bp	
*	56260	57028:	contig of 769 bp	in length
*	57029	57128:	gap of 100 bp	
*	57129	57851:	contig of 723 bp	in length
*	57852	57951:	gap of 100 bp	
*	57952	58690:	contig of 739 bp	in length
*	58691	58790:	gap of 100 bp	

	Query Match	8.7%; Score 48.2; DB 69; Length 217412;
	Best Local Similarity	48.6%; Pred. No. 1.5;
	Matches 160; Conservative	0; Mismatches 168; Indels 1; Gaps 1;
Oy	122 CATTAAAGCATCTGACACATGCCTGTATAAAGCAGTGGTCGCCATTAATCAATCGAGTCC	181
Dd	62840 CATGATATGATATACGAGATATCTTCATATATATATATGCTTGCCATATTTCAATGATATATA	62781
Oy	182 TTGCATTTGGGGGAAGAATAATGTTATAGTTTGTTTAGAAACTTTATTGCTGGAGATA	241
Dd	62780 TATTCATAGATAGATTCATATAATAGAGATGTGAATTAATTCATATATATATTCATATA	62721
Oy	242 GAGTTAAAGATCTTTAATCATATCCATGCAGGGGATWTCTTAAGTAATAGGGCTTTTCG	301
Dd	62720 TATTCATATATATATATATTCATATATATATATATATATATATATATATATATATATTC	62661
Oy	302 ACGAAGTATATGGTTGGATACAT-TTCCTAGTCGACGATATGTTTTTCTTATCAATATA	360
Dd	62660 ATAAATGATATATTCATTC	62601
Oy	361 GCATTTTCATAGCATATATATACATCTTGCACAATCATGAGATAGCTGCTACTATATA	420
Dd	62600 TATATATTCATATATATATATTCATGATAGTATATTCATATATATATATATATATATAT	62541
Oy	421 GGTCATTTGGCAATGGTATCTACACTT	449
Dd	62540 TCATATATATATATTCATATATATATATTCAT	62512
RESULT 2		
PFCOMP/RB/c	PECOMP/RB	14001 bp DNA INV 14-FEB-1997
LOCUS	Locus	P.falciparum complete gene map of plastid-like DNA (R-B).
DEFINITION	X95276	
ACCSSION	X95276.1 GI.1171591	
KEYWORDS	CIP gene; LSO rRNA gene; ORF105; ORF129; ORF78; ORF79; ORF91; rpII4 gene; rpII6 gene; rpII gene; rpI23 gene; rpI36 gene; rpI4 gene; rpI6 gene; rpsII gene; rpsI2 gene; rpsI7 gene; rpsI9 gene; rps3 gene; rps4 gene; RPS5 gene; RPS7 gene; rps7 gene; trnA-ASP; trnA-Cys; trnA-Gln; trnA-Glu; trnA-Gly; trnA-His; trnA-Ieu; trnA-Lys; trnA-Met; trnA-Phe; trnA-Pro; trnA-Ser; trnA-Thr; trnA-Trp; trnA-Tyr; tula gene.	
SOURCE	malaria parasite P. falciparum.	
ORGANISM	Plasmodium falciparum	
REFERENCE	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	
AUTHORS	Wilson,R.J.M., Denny,P.W., Preiser,P.R., Rangachari,K., Roberts,K., Roy,A., Whyte,A., Strach,M., Moore,D.J., Moore,P.W. and Williamson,G.H.	
TITLE	Complete gene map of the plastid-like DNA of the malaria parasite	

JOURNAL	Plasmodium falciparum
MEDLINE	J. Mol. Biol. 261 (2), 155-172 (1996)
REFERENCE	96346169
AUTHORS	2 (bases 1 to 14001)
TITLE	Wilson, R.J.M.
JOURNAL	Direct Submission
COMMENT	Submitted (22-JAN-1996) R.J.M. Wilson, National Institute for Medical Research, Mill Hill, London NW7 1AA, UK
FEATURES	Related sequences X90351-X90354 (incl.), and X87630-X87631 (incl.).
SOURCE	Location/Qualifiers
gene	1. .14001
gene	/organism="Plasmodium falciparum"
gene	/strain="C10"
gene	/db_xref="taxon:5833"
gene	/dev_stage="erythrocytic"
gene	/note="rR-B half of 35kb circle, putative plasmid DNA"
gene	<1. .227
gene	/gene="LSU rRNA"
gene	/product="large subunit ribosomal RNA"
gene	1. .227
gene	/gene="LSU rRNA"
gene	242. .314
gene	/gene="tRNA-Thr"
gene	/product="transfer RNA-Thr"
gene	242. .314
gene	/gene="tRNA-Thr"
gene	340. .966
gene	/gene="rps4"
gene	340. .966
gene	/gene="rps4"
gene	/codon_start=1
gene	/protein_id="CAA64576.1"
gene	/db_xref="GI:1171592"
gene	/db_xref="SPTREMBL:O25804"
gene	/translation="MIKELPKRIILKLNIPFLYSGRYKCLNKSYSKSPDIDKIRIRICYNVCITKYKRLLYLKNIDKNINILYELKLELRDLPLWNIEPFEETILORRYIKKKNYINNIINKYININKNNDIEFPNNKIKITILKLNILKYNNITYIISNLKYNIKIKIYSIKYKIFILICYNFKIKILNINILNLIYNDIYYI"
gene	981. .1053
gene	/gene="tRNA-His"
gene	/product="transfer RNA-His"
gene	981. .1053
gene	/gene="tRNA-His"
gene	1072. .1142
gene	/gene="tRNA-Cys"
gene	/product="transfer RNA-Cys"
gene	1072. .1142
gene	/gene="tRNA-Cys"
gene	1148. .1183
gene	/gene="tRNA-Leu"
gene	/note="extein 1"
gene	/product="transfer RNA-Leu"
gene	1148. .1183
gene	/gene="tRNA-Leu"
gene	1184. .1317
gene	/gene="tRNA-Leu"
gene	1184. .1317
gene	/gene="tRNA-Leu"
gene	1318. .1366
gene	/gene="tRNA-Leu"
gene	/note="extein 2"
gene	/product="transfer RNA-Leu"
gene	1318. .1366
gene	/gene="tRNA-Leu"
gene	1373. .1452
gene	/gene="tRNA-Met (1)"
gene	/product="transfer RNA-Met (1)"
gene	1373. .1452
gene	/gene="tRNA-Met (1)"
gene	1457. .1541
gene	/gene="tRNA-Tyr"
gene	/product="transfer RNA-Tyr"
gene	1457. .1541

gene
tRNA
1555..1646
/gene="tRNA-Tyr"
/gene="tRNA-Ser"
/product="transfer RNA-Ser"
1555..1646
/gene="tRNA-Ser"
1653..1728
/gene="tRNA-Asp"
/product="transfer RNA-asp"
1653..1728
/gene="tRNA-Asp"
1739..1811
/gene="tRNA-Lys"
/product="transfer RNA-Lys"
1739..1811
/gene="tRNA-Lys"
1819..1888
/gene="tRNA-Glu"
/product="transfer RNA-Glu"
1819..1888
/gene="tRNA-Glu"
1901..1972
/gene="tRNA-Pro"
/product="transfer RNA-Pro"
1901..1972
/gene="tRNA-Pro"
1999..2571
/gene="tp14"
1999..2571
/gene="tp14"
/codon_start=1
/protein_id="CAA64577.1"
/db_xref="GI:1171593"
/db_xref="SPTREMBL:O25805"
/translation="MTIIINNNLTNIIIFKKYNEFKLYENNYIKICKLIYYIKY
LYIYNIVAKHKHNSKSKYFESKRIKRVGKGAKRLKPKSPVKOGAGNGPPYKEN
KIITSINLRLEVLVLLINKRSNIITIKLENITNLNTIPYKNKCIFRKLAIKGINN
KYILINLNRLFNKNIFIINIIMNYLIPLI"
2575..2802
/gene="fp123"
2575..2802
/gene="fp123"
/codon_start=1
/protein_id="CAA64578.1"
/db_xref="GI:1171594"
/db_xref="SPTREMBL:O25806"
/translation="MKEVILNFYLNIIFFKIKINYLVNLPKCIYSISIKFYTKLDIKYIIKNN
IFKILIVNNINIKINNKKNCIAKKYITFEK"
2799..3536
/gene="fp12"
2799..3536
/gene="fp12"
/codon_start=1
/protein_id="CAA64579.1"
/db_xref="GI:1171595"
/db_xref="SPTREMBL:O25807"
/translation="MITLKIKYNTTYKIFKSGKANNKGYITTYNKGGNLKYNNKLIDS
WIDNNITINIKVLFKKIKNYFRNTYIGCLLYLSKLNNDQKFTIILOHNAMINISYLL
TINDNNIKGSYTOLKRGYICLVYINISDKYRKGSIFARSAGTFMOILSFYKNVIYL
PSKOCKYININFCYIGVANSNIFPNKFIRKAGYNIYNIKPVKGRKAHVGDHPHGCG
KGKGTIGRKRCPCSKKGLSHSGKYTIK"
3557..3829
/gene="tps19"
3557..3829
/gene="tps19"
/codon_start=1
/protein_id="CAA64580.1"
/db_xref="GI:1171596"
/db_xref="SPTREMBL:O25808"
/translation="MIKLYMLKATITNNKYIIRINNKRYKKNLNLNTYNNKLYIYKKL
LNLITYKVNIGKFLPIYITNKYTRLFNKLGNFITYTKVFANNIKELLIN"
3839..4483
/gene="tps3"

QY	454	TGAGGATTCGCGAGGCTGCTGCTATCCCGAGTATTAGCGGTATTCGTTTACCA	513
Db	192195	TTTTTTTTTAAATAATTTGTTTTTATTTTATTTTATTTCTAATAAATTTTTTTATATATATAAT	192136
QY	514	TTTCAT	519
Db	192135	TTTTTT	192130
RESULT	4	AC005139	256172 bp
LOCUS	AC005139/c	Plasmodium falciparum chromosome 12, *** SEQUENCING IN PROGRESS	HTG 01-APR-1999
DEFINITION		*** 5 unordered pieces.	
ACCESSION	AC005139		
VERSION	AC005139.3	GI:4558581	
KEYWORDS	HTG; HTGS; PHASE1.		
SOURCE		malaria parasite P. falciparum.	
ORGANISM		Plasmodium falciparum	
REFERENCE		Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	
AUTHORS		1 (bases 1 to 256172)	
TITLE		Hyman,R.W., Fung,E.L., Qin,F., Yamaki,T., Kurd,I.O.B. and Davys,R.W.	
REFERENCE		Plasmodium falciparum 3D7 chromosome 12	
JOURNAL		Unpublished	
AUTHORS		2 (bases 1 to 256172)	
TITLE		Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davys,R.W.	
JOURNAL		Direct Submission	
		Submitted (18-JUN-1998) Stanford DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA	
COMMENT		On Apr 2, 1999 this sequence version replaced gi:4337170.	
		* NOTE: This is a 'working draft' sequence. It currently	
		* consists of 5 contigs. The true order of the pieces	
		* is not known and their order in this sequence record is	
		* arbitrary. Gaps between the contigs are represented as	
		* runs of N, but the exact sizes of the gaps are unknown.	
		* This record will be updated with the finished sequence	
		* as soon as it is available and the accession number will	
		* be preserved.	
		1 126535: contig of 126535 bp in length	
		* 126536 126735: gap of unknown length	
		* 126736 158845: contig of 32110 bp in length	
		* 158846 159045: gap of unknown length	
		* 159046 211538: contig of 52493 bp in length	
		* 211539 211738: gap of unknown length	
		* 211739 248106: contig of 36368 bp in length	
		* 248107 248306: gap of unknown length	
		* 248307 256172: contig of 7666 bp in length.	
FEATURES		Location/Qualifiers	
Source		1..256172	
		/Organism="Plasmodium falciparum"	
		/db_xref="taxon:5833"	
		/chromosome="12"	
BASE COUNT	100700 a	26900 c	25826 g 101937 t 809 others
ORIGIN			
Query Match		8.6%; Score 47.6; DB 60; Length 256172;	
Best Local Similarity	43.6%;	Pred. No. 2;	
Matches	212; Conservative	0; Mismatches 274; Indels	0; Gaps 0;
QY	34	ATAAATATTATTCAGCTTCTTCCATTAACATATATGCATATCTCTATG	93
Db	145961	ATAAATTTTCCACACAGATGATTTGGTTTTTTCTTAAATAAATAACATATATGTATGTA	145902
QY	94	TGCGATCGTTTAATATATATGAAATATGATTAACGATTAACGATCTAGCATCTGTTAAAGC	153
Db	145901	TATATATGATTTAGTGATATATATTTTACGATATGATATGATGATTAATGATATATATATA	145842
QY	154	AGTGTTGCTCCAAATTTCAATCTAGCTCTTGCATTTGCTGGGACTAAATTTGTAATATTT	213

Db	145841	TATGTAATGTCTGAATCGTAAATGCTACTATTATGGCTATATATATATATATATATGTTAAATGTC	145782
Oy	214	GCTTTAAGAATACTTTATTCGTGAGTAGAGTTAAAGATCTTTAAATCATATCCATGGGT	273
Db	145781	AATGATTAATAAATTTATTTATTTATTTTTTTATTTATTTATTAATAAAAATTTTTTTATTTTTTTTTTTT	145722
Oy	274	CAATTTCTTAACCTAATTAATGGGGTTTTCACAGAACATATATGTTGCATACATTTCTACGTC	333
Db	145721	ATTATTAATTAATTTTTTTATTTTTTTTTTTTTTTTTTTTAAATTTATTTATTAATTTTTTTTTTTT	145667
Oy	334	TCAGATATTTGTTTTCTTATCAATATATAGCATAATTTCAATAGCTATATATCAATCTTGACAAA	393
Db	145661	ATTATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTT	145602
Oy	394	TCATATGATAGCTGTTCTTCTATATATAGTCAATATTCGATATGCGTATCTACACTTGGT	453
Db	145601	AACATTTTTTCCCCTTTCTTTCTTTTATTTATATATATATATATATTTTTTTTAAACATTTTTTTT	145542
Oy	454	TCAGATATTCGAGAGAGTGCTGCTATATCCAGATTATAGCGGATTCGTTTTTACCA	513
Db	145541	TTTTTATTTATTAATTAATTTGTTTTTATTTATTTATTTCTAATAAATTTTTTTTATATATTAAT	145482
Oy	514	TTCTATF	519
Db	145481	TTTTTT	145476
RESULT	5		
LOCUS	AC005140/c		
DEFINITION	Plasmodium falciparum chromosome 12 clone 3D7, ** SEQUENCING IN PROGRESS ***; 4 unordered pieces.		
ACCESSION	AC005140		
VERSION	AC005140.8		
KEYWORDS	HTG: HTGS.PHASE1.		
SOURCE	malaria parasite P. falciparum.		
ORGANISM	Plasmodium falciparum		
REFERENCE	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
AUTHORS	1 (bases 1 to 310779)		
JOURNAL	Hyman,R.W., Fung,E.L., Qin,F., Rowley,D., Mao,J., Tamaki,T., Kuri,O.B., Conway,A.B. and Davis,R.W.		
REFERENCE	Plasmodium falciparum 3D7 chromosome 12		
AUTHORS	Unpublished		
JOURNAL	2 (bases 1 to 310779)		
REFERENCE	Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W.		
AUTHORS	Direct Submission		
JOURNAL	Submitted (18-JUN-1998) Stanford DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA		
COMMENT	On Aug 12, 2000 this sequence version replaced gi:8810455. ** NOTE: This is a "working draft" sequence. It currently * consists of 4 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved. * * 1 187308: contig of 187308 bp in length * 187309 187508: gap of unknown length * 187509 257820: contig of 70312 bp in length * 257821 258020: gap of unknown length * 258021 307255: contig of 49225 bp in length * 307256 307455: gap of unknown length * 307456 310779: contig of 3324 bp in length. Location/Qualifiers 1. .310779 /organism="Plasmodium falciparum" /db_xref="taxon:5833" /chromosome="12" /clone="PFYCB8-420" /clone="3D7"		
FEATURES			
SOURCE			

SEYENDINDIYVPGSPKRYTLEEVLEPGSNNTASGKNTPSDTQNDIONDGPSTP
PNTSPDIPSPDIPPTIDDEWNTLTKDEFISOLQSEOPNDVNDKSGNSSTNTI
TTTSRANVEKPEPIMSJHBRDLTYGEEYDMSNSGIPSSNRDLSGTKYVSGI
DLJNDSNNGNOPIDIDEYELIKREKENELFGTNHYKOTSISVAKLINSDIOLNLEF
HKMLDRHRCCEKLNENRKLKEMEENHTSGNHTSDSNKTLNTDVSIOIDMD
GKPKRENTMDTLEDKDKNEPYVDODDIYDVNDHDVSTVDSNMDIPSKVQIEM
DVNTKLYKEKPYIADVMDT
38960..38965
/note="corrected splice donor for PFC0005c"
42416..43015
/note="R-PA3"
complement(join(43045..43971,44202..44255))
/gene="PFC0010c, MAL3P8.2"
complement(join(43045..43971,44202..44255))
/gene="PFC0010c, MAL3P8.2"
/note="PFC0010c (MAL3P8.2), r1fin, (3D7-r1fT3-1), len: 326
aa: Similarity to other P.falciptarum r1fins. Member of a
multi-gene family; predicted using hexonon"
/codon_start=1
/protein_id="CAB39116.1"
/db_xref="GI:4494021"
/db_xref="SPTREMBL:O97325"
/translation="MKIHYINILFLPLNLYNORNHITRPRKATRTICECELY
APSNVSDPEMKRVMOQFVDRTORFHEVDERMKTPRCKECCDKEIKILKDKLE
KOMAOOLTTTDPNITTEDITPTCEKESLAKTERFCINCRTMGVAPRGVLGCLCY
VGTNVTYOTDALOKGIEAGYKAGIEBKGPGLKILKVSQIOSFINPANYAEETTF
NFKOVENTCVKKEHTEFCYSTYISKENALSORSEIATKANADMAKLAEEVLDKAP
VTSGLNTAIIASIVAILVLMIIYILFLRYRKNKKKKKLYIKLEE"
complement(46448..47828)
/gene="PFC0015c"
complement(46448..47828)
/gene="PFC0015c"
/note="PFC0015c varc pseudogene"
/codon_start=1
/pseudo
complement(join(48956..49777,49881..49949))
/gene="PFC0025c, MAL3P8.4"
complement(join(48956..49777,49881..49949))
/gene="PFC0025c, MAL3P8.4"
/note="PFC0025c (MAL3P8.4), stevor (3D7-stevoT3-1), len:
296 aa: Similarity to other P.falciptarum stevors. Member
of a multi-gene family; predicted using hexonon"
/codon_start=1
/protein_id="CAB39117.1"
/db_xref="GI:4494022"
/translation="MKMYLYKMLLFTFLINTLIPHYENLYNDYNSFYONTKRTT
IKSRLLAOTQIHNPHYHNDPELKEIIDKMEBAIKRYOQHDPYEOLEKVEYKNGTRY
SGCNDAPETLLEKELLEYEELFEGNSDMLKSGMSVNDKESSTPECCTDINGVIAK
TKGRDYIKHLKGRGTGYFCGLSGVILITGTAAKGAISTPAGSGKNCIYATSI
FHLNKKSSSATFALCTGVCADYASGATSAATFPCIAALVLLITLVLLITIT
WLRRKRSWKHECKHCK"
complement(join(52066..53076,53192..53260))
/gene="PFC0030c, MAL3P8.5"
complement(join(52066..53076,53192..53260))
/gene="PFC0030c, MAL3P8.5"
/note="PFC0030c (MAL3P8.5), r1fin, (3D7-r1fT3-2), len: 359
aa: Similarity to other P.falciptarum r1fins. Member of a
multi-gene family; predicted using hexonon"
/codon_start=1
/protein_id="CAB39118.1"
/db_xref="GI:4494023"
/db_xref="SPTREMBL:O97326"
/translation="MKIYNSNILLFALANLILTYANKNKPSITPHHTRTTSL
LSCDRESSIYDSEINYSKEIFEROTSRFEYERMTTORAKKEODKIOIKI
HKMKENLNALEKIEKGLRCGGLGASGIGFVAVIMRPAALDAIAKATIEG
TANISAGVAGATGKVLIVISGLKMGISTGKNLSEYFATTSYKKAIVSAVVE
OHEKCEPGYKGPVGVDSRDIHRCOSWMOGTSVSTGYISKVDIKRTVVE
VSAABPANAARFVATERTATIKABEKITEASIOULSAIGSIALIILVIMLT
ITILYRKKKKKKKAAQYTKLINE"
join(58103..58174,58338..59390)
/gene="PFC0035w, MAL3P8.6"
join(58103..58174,58338..59390)
/gene="PFC0035w, MAL3P8.6"

Query Match 8.58; Score 47.2; DB 6; Length 108908;
Best Local Similarity 47.98; Pred. No. 2.5;
Matches 136; Conservative 0; Mismatches 148; Indels 0; Gaps 0;
/note="PFC0035w (MAL3P8.6), r1fin, (3D7-r1fT3-3), len: 375
aa: Similarity to other P.falciptarum r1fins. Member of a
multi-gene family; predicted using hexonon"
/codon_start=1
/protein_id="CAB39119.1"
/db_xref="GI:4494024"
/db_xref="SPTREMBL:O97327"
/translation="MRITIKMKVYINILLLALSNILANTHOKSPSPRHIOITRL
CECELYHNYNDNPEKRVMOQFDRTORFHEVDERMKTPRCKECCDKEIKILKDKLE
KDKTEKLNKFPATLORDIHSDAIPDTICSTNADKREKCTICNTGIVAPSSG
LAGIABEALYAMKRNALQTLIENALKAANDILVGGIEAKKYYIGSLDLTEKIGI
GSEWPTFEGICINVSASTIIEKROTLGATSTIDKATVCQISIGTQWHDGTY
GAGPTPIETVLNGIVEGTKEADVAEAARESATNAIKVQETRLLEAGNSSISSIN
ASIIAIVILIIWIIYILFLRYRKNKKKKKLYIKLEE"
join(61248..61301,61548..62459)
/gene="PFC0040w, MAL3P8.7"
join(61248..61301,61548..62459)
/gene="PFC0040w, MAL3P8.7"
/note="PFC0040w (MAL3P8.7), r1fin, (3D7-r1fT3-4), len: 322
aa: Similarity to other P.falciptarum r1fins. Member of a
multi-gene family; predicted using hexonon"
/codon_start=1
/protein_id="CAB39120.1"
/db_xref="GI:4494025"
/db_xref="SPTREMBL:O97328"
/translation="MKIHYINILFLPLNLYNORNHITRPRKATRTICECELY
APSNVSDPEMKRVMOQFVDRTORFHEVDERMKTPRCKECCDKEIKILKDKLE
KOMAOOLTTTDPNITTEDITPTCEKESLAKTERFCINCRTMGVAPRGVLGCLCY
VGTNVTYOTDALOKGIEAGYKAGIEBKGPGLKILKVSQIOSFINPANYAEETTF
NFKOVENTCVKKEHTEFCYSTYISKENALSORSEIATKANADMAKLAEEVLDKAP
VTSGLNTAIIASIVAILVLMIIYILFLRYRKNKKKKKLYIKLEE"
complement(46448..47828)
/gene="PFC0015c"
complement(46448..47828)
/gene="PFC0015c"
/note="PFC0015c varc pseudogene"
/codon_start=1
/pseudo
complement(join(48956..49777,49881..49949))
/gene="PFC0025c, MAL3P8.4"
complement(join(48956..49777,49881..49949))
/gene="PFC0025c, MAL3P8.4"
/note="PFC0025c (MAL3P8.4), stevor (3D7-stevoT3-1), len:
296 aa: Similarity to other P.falciptarum stevors. Member
of a multi-gene family; predicted using hexonon"
/codon_start=1
/protein_id="CAB39117.1"
/db_xref="GI:4494022"
/translation="MKMYLYKMLLFTFLINTLIPHYENLYNDYNSFYONTKRTT
IKSRLLAOTQIHNPHYHNDPELKEIIDKMEBAIKRYOQHDPYEOLEKVEYKNGTRY
SGCNDAPETLLEKELLEYEELFEGNSDMLKSGMSVNDKESSTPECCTDINGVIAK
TKGRDYIKHLKGRGTGYFCGLSGVILITGTAAKGAISTPAGSGKNCIYATSI
FHLNKKSSSATFALCTGVCADYASGATSAATFPCIAALVLLITLVLLITIT
WLRRKRSWKHECKHCK"
complement(join(52066..53076,53192..53260))
/gene="PFC0030c, MAL3P8.5"
complement(join(52066..53076,53192..53260))
/gene="PFC0030c, MAL3P8.5"
/note="PFC0030c (MAL3P8.5), r1fin, (3D7-r1fT3-2), len: 359
aa: Similarity to other P.falciptarum r1fins. Member of a
multi-gene family; predicted using hexonon"
/codon_start=1
/protein_id="CAB39118.1"
/db_xref="GI:4494023"
/db_xref="SPTREMBL:O97326"
/translation="MKIYNSNILLFALANLILTYANKNKPSITPHHTRTTSL
LSCDRESSIYDSEINYSKEIFEROTSRFEYERMTTORAKKEODKIOIKI
HKMKENLNALEKIEKGLRCGGLGASGIGFVAVIMRPAALDAIAKATIEG
TANISAGVAGATGKVLIVISGLKMGISTGKNLSEYFATTSYKKAIVSAVVE
OHEKCEPGYKGPVGVDSRDIHRCOSWMOGTSVSTGYISKVDIKRTVVE
VSAABPANAARFVATERTATIKABEKITEASIOULSAIGSIALIILVIMLT
ITILYRKKKKKKKAAQYTKLINE"
join(58103..58174,58338..59390)
/gene="PFC0035w, MAL3P8.6"
join(58103..58174,58338..59390)
/gene="PFC0035w, MAL3P8.6"

Man, K. H., Doyle, C., Baxter, E. G., Helt, G., Nelson, C. R., Gabor
 Miklos, G. L., Abtil, J. F., Agbayan, A., An, H. J.,
 Andrews-Pfannkoch, C., Baldwin, D., Ballew, R. M., Basu, A.,
 Bensadoun, J., Bhatnagar, L., Beasley, E. M., Beeson, K. Y.,
 Bencos, P. V., Berman, B. P., Bhandari, D., Bolshakov, S., Borkova, D.,
 Botchan, M. R., Bouck, J., Brokstein, P., Brotlier, P., Burtis, K. C.,
 Busam, D. A., Butler, H., Cadieu, E., Center, A., Chandra, I.,
 Cherry, J. M., Cawley, S., Dahlke, C., Davenport, L. B., Davies, P., de
 Pablo, B., Delcher, A., Deng, Z., Mays, A. D., Dev, I., Dietz, S. M.,
 Dodson, K., Dou, L. E., Downes, M., Dugan-Rocha, S., Dunkov, B. C.,
 Dunn, P., Durbin, K. J., Evangelista, C. C., Ferraz, C., Ferriera, S.,
 Fleischmann, W., Foster, C., Gabrielian, A. E., Garg, N. S.,
 Gelbart, W. M., Glasser, K., Glodek, A., Gong, F., Gorrell, J. H., Gu, Z.,
 Guan, P., Harris, M., Harris, N. L., Harvey, D., Heiman, T. J.,
 Hernandez, J. R., Houck, J., Hostin, D., Houston, K. A., Howland, T. J.,
 Wei, M. H., Iobagiam, C., Jalali, M., Kalush, F., Karpene, G. H., Ke, Z.,
 Kenison, J. A., Ketchum, K. A., Kimmel, B. E., Kodira, C. D., Kraft, C.,
 Kravitz, S., Kulp, D., Lai, Z., Lasko, P., Lei, Y., Levitsky, A. A.,
 Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Mattei, B., McIntosh, T. C.,
 McLeod, M. P., McPherson, D., Merkulov, G., Mitsuhashi, N. V., Mobarry, C.,
 Morris, J. A., Moshrefi, A., Mount, S. M., Moy, M., Murphy, B., Murphy, L.,
 Muzny, D. M., Nelson, D. L., Nelson, D. R., Nelson, K. A., Nixon, K.,
 Nusskern, D. R., Pacle, J. M., Palazzolo, M., Pittman, G. S., Pan, S.,
 Pollard, J., Puri, V., Reese, M. G., Reinert, K., Remington, K.,
 Saunders, R. D., Scheeler, F., Shen, H., Shue, B. C., Siden-Kiamos, I.,
 Simpson, M., Skupski, M. P., Smith, T., Spier, E., Spradling, A. C.,
 Stapleton, M., Strong, R., Sun, E., Svirska, R., Tector, C., Turner, R.,
 Venter, E., Wang, A. H., Wang, X., Wang, Z. Y., Wasserman, D. A.,
 Weinstock, G. M., Weissbach, J., Williams, S. M., Woodage, T.,
 Wolley, K. C., Wu, D., Yang, S., Yao, Q. A., Ye, J., Yeh, R. F.,
 Zaveri, J. S., Zhang, M., Zhang, G., Zhao, Q., Zheng, X. H.,
 Zhong, F. N., Zhong, W., Zhou, S., Zhu, X., Zhu, X., Smith, H. O.,
 Gibbs, R. A., Myers, E. W., Rubin, G. M., and Venter, J. C.
 The genome sequence of *Drosophila melanogaster*
 Science 287 (5461), 2185-2195 (2000)

2 (bases 1 to 298469)
 20196006

JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 JOURNAL
 COMMENT
 FEATURES
 SOURCE

Adams, M. D., Celisner, S. E., Gibbs, R. A., Rubin, G. M. and Venter, C. J.
 Direct Submission
 Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
 Rockville, MD, USA
 On Oct 9, 2000 this sequence version replaced gi:7304344.
 Location/Qualifiers
 1..298469

/organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /chromosome="4"
 complement(1028..>1306)
 /gene="CG17469"
 /product="CT38623"
 /db_xref="FLYBASE:FBan0017469"
 /db_xref="FLYBASE:FBgn0040845"
 complement(<1028..>1306)
 /gene="CG17469"
 /map="102E1-102E3"
 /db_xref="FLYBASE:FBan0017469"
 /db_xref="FLYBASE:FBgn0040845"
 complement(1028..1192)
 /gene="CG17469"
 /note="CG17469 gene product"
 /codon_start=1
 /db_xref="FLYBASE:FBan0017469"
 /db_xref="FLYBASE:FBgn0040845"
 /protein_id="AAF59375.1"
 /db_xref="GI:7304345"
 /translation="MTESGIDLGDFMEFDLNLINLNDNDNDPLPNTENNEYELKS
 SSRICIRHEW"
 join(3595..3608,3730..3887,3945..4042,4116..4365,
 4417..4547,4611..4779,6348..6499,7372..7501,7555..>7671)
 /gene="CG17471"
 /product="CT38625"
 /db_xref="FLYBASE:FBan0017471"
 /db_xref="FLYBASE:FBgn0039924"

mrna
 cds
 gene

gene
 cds
 mrna
 gene
 cds
 gene
 mrna
 cds
 gene

<3595..>7671
 /gene="CG17471"
 /map="102E1-102E3"
 /db_xref="FLYBASE:FBan0017471"
 /db_xref="FLYBASE:FBgn0039924"
 join(3750..3887,3945..4042,4116..4365,4417..4547,
 4611..4779,6348..6499,7372..7501,7555..7671)
 /gene="CG17471"
 /note="CG17471 gene product"
 /codon_start=1
 /db_xref="FLYBASE:FBan0017471"
 /db_xref="FLYBASE:FBgn0039924"
 /protein_id="AAF59376.2"
 /db_xref="GI:10728139"
 /translation="MEKISSSSOPRLKKKHPYKXKVKLFRANEPILSVFMGICN
 HTNELSGVNPVWLPPDPRAYSKIKYDNLFPKKNPSPFKYKRYEPLFRNLRER
 FGVDVDFRESLTISOPLOIDSSGSGKQFQSTDKFTITSLTSEETERNHAKQY
 HPYVERGRKTLFQYLGMRYITVESGYVYFVKNRWFSSHLTHKFKDLGSDVRE
 ASEKELEKNLPTFKDNDIFKOKYKIDIKERAKDKIMDTLNDVLTLEHIMDSILV
 GHDCVRAEEALOGDNLITVGRSENSEECDSGERATVTGKIKELMYILNCIA
 EKRELYFAIIDVLTQYGVKKQAKAATVAGSVNDISTCDEPQAKRRLDPMKRA
 IE"
 complement(join(7685..7850,7901..7982,8045..8187,
 9658..10140,10199..10336))
 /gene="CG17467"
 /product="CT38621"
 /db_xref="FLYBASE:FBan0017467"
 /db_xref="FLYBASE:FBgn0022361"
 complement(<7685..>10336)
 /gene="CG17467"
 /map="102E1-102E3"
 /db_xref="FLYBASE:FBan0017467"
 /db_xref="FLYBASE:FBgn0022361"
 complement(join(7848..7850,7901..7982,8045..8187,
 9658..10140,10199..10337))
 /gene="CG17467"
 /note="CG17467 gene product"
 /codon_start=1
 /db_xref="FLYBASE:FBan0017467"
 /db_xref="FLYBASE:FBgn0022361"
 /protein_id="AAF59377.1"
 /db_xref="GI:7304347"
 /translation="MKLDANIKSGLSNHVQYTLDPKKYQIHNOVKQENSDEIVKSDS
 KPEEVKATIELSKKFEFFYSPEAKNFRFTPTPEGVNFPPEPSEINOOISIT
 LFNNILOKTKTSTILNKFERMEKQKDDOKKPNPLCTFPPPIISOPISRTD
 EESDSDEKNSDENSEDEINPSNSYDKALKLAQSAKQIKRAKEREKQNNIEIO
 EIKGRIDVYSQSLSNESIESAKAIRERFERMKRSITAMENPSQTQIKRFY"
 join(18584..18839,18906..19163,25083..25226,25280..26049,
 26111..26402,27005..27124,27178..27740)
 /gene="pho"
 /note="Nucleotide sequence of the Celera sequence differs
 from the published sequence for this transcript."
 /product="CT39329"
 /db_xref="FLYBASE:FBan0017743"
 /db_xref="FLYBASE:FBgn0002521"
 <18584..>27740
 /gene="pho"
 /note="CG17743"
 /map="102E3-102E5"
 /db_xref="FLYBASE:FBan0017743"
 /db_xref="FLYBASE:FBgn0002521"
 join(19069..19163,25083..25226,25280..26049,26111..26402,
 27005..27124,27178..27325)
 /gene="pho"
 /note="pho gene product; Nucleotide sequence of the Celera
 sequence differs from the published sequence for this
 transcript"
 /codon_start=1
 /db_xref="FLYBASE:FBan0017743"
 /db_xref="FLYBASE:FBgn0002521"
 /protein_id="AAF59378.1"
 /db_xref="GI:7304348"
 /translation="MAYERFGIILIQSEYDEDIGNTRYVKNKMGKGNHYDLHRKNAP"

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 193742)
TITLE	Burton,J
JOURNAL	Direct Submission
COMMENT	Submitted (09-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Feb 12, 2001 this sequence version replaced gl:12743866.

[illegible]

Query Match	8.1%	Score 45.2	DB 6	Length 3171
Best Local Similarity	45.5%	Pred. No. 8.7		
Matches 161	Conservative 0	Mismatches 193	Indels 0	Gaps 0
BASE COUNT	1538 a	502 c	248 g	883 t
ORIGIN				
Query Match	8.1% <td>Score 45.2 <td>DB 6 <td>Length 3171</td> </td></td>	Score 45.2 <td>DB 6 <td>Length 3171</td> </td>	DB 6 <td>Length 3171</td>	Length 3171
Best Local Similarity	45.5% <td>Pred. No. 8.7</td> <td></td> <td></td>	Pred. No. 8.7		
Matches 161	Conservative 0	Mismatches 193	Indels 0	Gaps 0
BASE COUNT	1538 a	502 c	248 g	883 t
ORIGIN				
Query Match	8.1% <td>Score 45.2 <td>DB 6 <td>Length 3171</td> </td></td>	Score 45.2 <td>DB 6 <td>Length 3171</td> </td>	DB 6 <td>Length 3171</td>	Length 3171
Best Local Similarity	45.5% <td>Pred. No. 8.7</td> <td></td> <td></td>	Pred. No. 8.7		
Matches 161	Conservative 0	Mismatches 193	Indels 0	Gaps 0
BASE COUNT	1538 a	502 c	248 g	883 t
ORIGIN				

9419 9518: gap of 100 bp
9519 10521: contig of 1003 bp in length
10522 10621: gap of 100 bp
10622 11585: contig of 964 bp in length
11586 11685: gap of 100 bp
11686 12674: contig of 989 bp in length
12675 12774: gap of 100 bp
12775 13775: contig of 1001 bp in length
13776 13875: gap of 100 bp
13876 14783: contig of 908 bp in length
14784 14883: gap of 100 bp
14884 15865: contig of 982 bp in length
15866 15965: gap of 100 bp
15966 16944: contig of 979 bp in length
16945 17044: gap of 100 bp
17045 17985: contig of 941 bp in length
17986 18085: gap of 100 bp
18086 19048: contig of 963 bp in length
19049 19148: gap of 100 bp
19149 20106: contig of 958 bp in length
20107 20206: gap of 100 bp
20207 21176: contig of 970 bp in length
21177 21276: gap of 100 bp
21277 22256: contig of 980 bp in length
22257 22356: gap of 100 bp
22357 23336: contig of 980 bp in length
23337 23436: gap of 100 bp
23437 24405: contig of 969 bp in length
24406 24505: gap of 100 bp
24506 25457: contig of 952 bp in length
25458 25557: gap of 100 bp
25558 26541: contig of 964 bp in length
26542 26641: gap of 100 bp
26642 27615: contig of 974 bp in length
27616 27715: gap of 100 bp
27716 28695: contig of 980 bp in length
28696 28795: gap of 100 bp
28796 29720: contig of 925 bp in length
29721 29820: gap of 100 bp
29821 30786: contig of 966 bp in length
30787 30886: gap of 100 bp
30887 31838: contig of 952 bp in length
31839 31938: gap of 100 bp
31939 32885: contig of 947 bp in length
32886 32985: gap of 100 bp
32986 33960: contig of 975 bp in length
33961 34060: gap of 100 bp
34061 35062: contig of 1002 bp in length
35063 35162: gap of 100 bp
35163 36155: contig of 997 bp in length
36160 36259: gap of 100 bp
36260 37224: contig of 965 bp in length
37225 37324: gap of 100 bp
37325 38293: contig of 969 bp in length
38294 38393: gap of 100 bp
38393 39377: contig of 984 bp in length
39378 39477: gap of 100 bp
39478 40410: contig of 933 bp in length
40411 40510: gap of 100 bp
40511 41478: contig of 968 bp in length
41479 41578: gap of 100 bp
41579 42553: contig of 975 bp in length
42554 42653: gap of 100 bp
42654 43612: contig of 959 bp in length
43613 43712: gap of 100 bp
43713 44689: contig of 977 bp in length
44690 44789: gap of 100 bp
44790 45718: contig of 929 bp in length
45719 45818: gap of 100 bp
45819 46747: contig of 929 bp in length
46748 46847: gap of 100 bp
46848 47815: contig of 968 bp in length
47816 47915: gap of 100 bp

47916 48930: contig of 1015 bp in length
48931 49030: gap of 100 bp
49031 50005: contig of 975 bp in length
50006 50105: gap of 100 bp
50106 51093: contig of 998 bp in length
51094 51193: gap of 100 bp
51194 52181: contig of 988 bp in length
52182 52281: gap of 100 bp
52282 53272: contig of 991 bp in length
53273 53372: gap of 100 bp
53373 54370: contig of 998 bp in length
54371 54470: gap of 100 bp
54471 55469: contig of 999 bp in length
55470 55569: gap of 100 bp
55570 56544: contig of 975 bp in length
56545 56644: gap of 100 bp
56645 57668: contig of 1024 bp in length
57669 57768: gap of 100 bp
57769 58728: contig of 960 bp in length
58729 58828: gap of 100 bp
58829 59829: contig of 1001 bp in length
59830 59929: gap of 100 bp
59930 60913: contig of 984 bp in length
60914 61013: gap of 100 bp
61014 62013: contig of 1000 bp in length
62014 62113: gap of 100 bp
62114 63077: contig of 964 bp in length
63078 63177: gap of 100 bp
63178 64116: contig of 939 bp in length
64117 64216: gap of 100 bp
64217 65161: contig of 945 bp in length
65162 65261: gap of 100 bp
65262 66233: contig of 972 bp in length
66234 66333: gap of 100 bp
66334 67313: contig of 980 bp in length
67314 67413: gap of 100 bp
67414 68394: contig of 981 bp in length
68395 68494: gap of 100 bp
68495 69484: contig of 990 bp in length
69485 69584: gap of 100 bp
69585 70506: contig of 922 bp in length
70507 70606: gap of 100 bp
70607 71547: contig of 941 bp in length
71548 71647: gap of 100 bp
71648 72629: contig of 982 bp in length
72630 72729: gap of 100 bp
72730 73654: contig of 925 bp in length
73655 73754: gap of 100 bp
73755 74680: contig of 926 bp in length
74681 74780: gap of 100 bp
74781 75782: contig of 1002 bp in length
75783 75882: gap of 100 bp

Query Match 8.18; Score 44.8; DB 67; Length 84346;
Best Local Similarity 46.28; Pred. No. 7.9;
Matches 148; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

OY 110 TATGGAATATGCAATTAAGCATGATGACATGCTGAATAAGAGAGTTCCTCAATAT 169
Db 61163 TGTATAGCACTTCTTCTTCTGTAAGCTTGCAGTTATATGTTTATGACTATTT 61222
OY 170 TCAATCTAGGCTTGCATTTGGTGAGATACATTAATGTTATGTTGTTAAGAAATCTTT 229
Db 61223 AAAAATATGATCCGACCTTGTAAGATGATATTCATGTTGTTGATTTACATTT 61282
OY 230 ATTGCTGAGTAGAGTTAAAGAGCTTTAATCATATATCCATGGGTATTTTAACTTAA 289
Db 61283 CTTGATGTTTATGATGATATGATGATGATTTTTCATGTTGTTGCTGCTATCCCTCT 61342
OY 290 TAGGAGTTTTCAGAGAAATATATGTTGATGATATTTCTGATGCTGATGATTTTCT 349
Db 61343 TTGAGAGATATCTCCGATGATGTTATTTGATGCTTTTATATAGAAATATATTTTCT 61402


```

CDS
/gene="ATP8"
4115..4261
/gene="ATP8"
/codon_start=1
/transl_table=4
/evidence=not_experimental
/product="ATPase subunit 8"
/protein_id="BA06565.2"
/db_xref="GI:11036339"
/translacion="MPQLVPEYFLNQLTYGFLILLIVLFSQFLPRILRLYSRLF
ISRL"
4441..5211
/gene="ATP6"
4441..5211
/gene="ATP6"
/codon_start=1
/transl_table=4
/evidence=not_experimental
/product="ATPase subunit 6"
/protein_id="BA06566.2"
/db_xref="GI:11036340"
/translacion="MTNIVNSPLDQFDIKYIGFVSPEIDININITTFVSEVIL
IVILGLDNDNRITIGNRMYVSQAMYDINMVGSGIGKIGYFFPLTYFFIF
IFTNLIGMIPYSRAITSQHWYFITSISVIMLCVTIIGLTHGILTPALFVPACPIA
LAPLVLELISARAIISGLRSANLSCHLMVLDGLVFNMSVSTFVLGFI
PLAGILVALEFAIMIQSVFAILASGYIKDLYLH"
5299..5369
/gene="tRNA-I"
/note="anticodon-gau"
/evidence=not_experimental
/product="transfer RNA-Ile"
5299..5369
/gene="tRNA-I"
5460..5546
/gene="tRNA-SI"
/note="anticodon-uga"
/evidence=not_experimental
/product="transfer RNA-Ser1"
5460..5546
/gene="tRNA-SI"
5610..5680
/gene="tRNA-A"
/note="anticodon-ugc"
/evidence=not_experimental
/product="transfer RNA-Ala"
5610..5680
/gene="tRNA-A"
5803..5889
/gene="tRNA-y"
/note="anticodon-gua"
/evidence=not_experimental
/product="transfer RNA-Tyr"
5803..5889
/gene="tRNA-y"
6034..6107
/gene="tRNA-N"
/note="anticodon-guu"
/evidence=not_experimental
/product="transfer RNA-Asn"
6034..6107
/gene="tRNA-N"
6330..6953
/gene="ND6"
6330..6953
/gene="ND6"
6330..6953
CDS
Query Match 8.0% Score 44.4 DB 14 Length 27694;
Best Local Similarity 44.6% Pred. No. 10;
Matches 174; Conservative 0; Mismatches 216; Indels 0; Gaps 0;
QY 42 TATTCAGCTTTTCCATTAACATGATGCATTAATCTATGTCGTCATG 101
DB 23418 TAATCCACCTTATGTTTCCTCCATATATTAATATCATTTATTAATGTAATTATC 23359

```

```

QY 102 GTTAAATATATGGAATATATGCAATTAAGCAGATCTAGACATGCTGTTAAAGCAGTGTGC 161
DB 23358 ATTAAATAGATATATTAATATGATTTGATTCATATATTAATTAATTAATTAATTAAT 23299
QY 162 TCCAAATATTCAGTACGTCTTGCAATGTCGGGATTAATGTTAGTGGTTGTTAG 221
DB 23298 TAAATAAATTTGATATATATATCTTTATTTATATATTAATTAATTAATTAATTAATTAAT 23239
QY 222 AAATCTTATTCGTCGAGTATAGATTAAGATCTTAATCAATTCAGGATGATTAATCT 281
DB 23238 TAATATATTAATTTATATATATTAATTTATTTGATGATTAATTAATTAATTAATTAATTAAT 23179
QY 282 TAACCTTAATAGGGGTTTTCGACAGATTAATGTTGATTAATTAATTAATTAATTAATTAAT 341
DB 23178 ATCATTTATAGTATATTAATTTGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 23119
QY 342 GTTTTCTTATCAATTAATAGCATATTTGATTAATTAATTAATTAATTAATTAATTAATTAAT 401
DB 23118 TATGTATAGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 23059
QY 402 AGCTGTCTACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 431
DB 23058 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 23029

```

```

RESULT 14
AC027632
LOCUS
DEFINITION
Homo sapiens chromosome 1 clone RP11-487E1, WORKING DRAFT SEQUENCE,
23 unordered pieces.
ACCESSION
AC027632
VERSION
AC027632.6 GI:9929791
KEYWORDS
HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 185926)
AUTHORS
Waterston,R.H.
TITLE
The sequence of Homo sapiens clone
JOURNAL
unpublished
AUTHORS
Waterston,R.H.
REFERENCE
2 (bases 1 to 185926)
JOURNAL
Direct Submision
AUTHORS
Submitted (30-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT
On Aug 26, 2000 this sequence version replaced gi:9838225.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Project Information
Center project name: H_NH0487E01
Summary Statistics
Sequencing vector: M13; 95%
Sequencing vector: plasmid; 5%
Chemistry: Dye-terminator Big Dye; 93% of reads
Chemistry: Dye-terminator Big Dye; 7% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 172750 bases at least Q40
Consensus quality: 176629 bases at least Q30
Consensus quality: 178424 bases at least Q20
Insert size: 191000; agarose-fp
Insert size: 184940; sum-of-contrigs
Quality coverage: 4.41 in Q20 bases; agarose-fp
Quality coverage: 4.66 in Q20 bases; sum-of-contrigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 23 contigs. The true order of the pieces
* is not known and their order in this sequence record is

```



```

REFERENCE 1 (bases 1 to 118313)
AUTHORS Murray,J., Rohlfing,T. and Antoniou,B.
TITLE The sequence of H. sapiens BAC clone CTB-167B5
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 118313)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-1998) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 3 (bases 1 to 118313)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 4 (bases 1 to 118313)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (04-FEB-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT ----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_RG167B05
-----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate
chemistry; an attempt was made to resolve all sequencing problems,
such as compressions and repeats; all regions were covered by
sequence from more than one subclone; and the assembly was
confirmed by restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and
sequencing collaboration between the NHGRI Chromosome 7 Mapping
Project (Eric D. Green, Director), John D. McPherson in the
Department of Genetics (Washington University), and the Washington
University Genome Sequencing Center. For additional information
about the map position of this sequence, see
http://www.nhgri.nih.gov/DIR/CTB/CHR7 or send
mailto:egreen@nhgri.nih.gov

SOURCE INFORMATION:
Clone CTB-167B5 is from the first release of the human BAC library
CTB-978SK-B. The library contains cloned DNA from the male
fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl.
Acad. Sci. USA 89:8794-7 (1992); U.-J. Kim et al., Genomics 34:213-8
(1996). This clone is available from Research Genetics, Inc.
(http://www.resgen.com).
VECTOR: pBelorAC11
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:
The actual start of this clone is at base position 1 of CTB-167B5;
actual end of this clone is at base position 118313 of CTB-167B5.

This clone contains SRS SWS52908 (NID:91113638).
Location/Qualifiers
1. 118313
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/clone="CTB-167B5"
/clone_1lb="CTB-978SK-B"
/map="7q21"

repeat_region 353. 519
/rpt_family="MER1_type"
repeat_region 527. 626
/rpt_family="MER1_type"
repeat_region 686. 832
/rpt_family="MER1_type"
repeat_region 914. 999
/rpt_family="L2"
repeat_region 1073. 1117
/rpt_family="L2"
repeat_region 1359. 1503
/rpt_family="L2"
repeat_region 3917. 4205
/rpt_family="Alu"
repeat_region 4701. 4865
/rpt_family="Alu"
gene complement(<5050. 7954)
/ gene="SRI"
CDS complement({join(<5050. 5119,6787. 6870,7904. 7954)})
/ gene="SRI"
/ note="match to sorcin protein P30626 (PID:9267021);
H_RG167B05.1"
/codon_start=1
/product="calcium binding protein amplified in
mutidrug-resistant cells"
/protein_id="AAB96327.1"
/db_xref="GI:2772536"
/translacion="MAYPGHAGAGGCGYPPGCGAGPAGPAGPAGTODPLXGFAAAG
ODGOLDADDERORCLTOSGINGYK"
9625. 10516
/rpt_family="L1"
repeat_region 10518. 11038
/rpt_family="L1"
repeat_region 11200. 11288
/rpt_family="MIR"
repeat_region 11353. 11404
/rpt_family="MER1_type"
repeat_region 11466. 11711
/rpt_family="MER1_type"
repeat_region 12803. 12913
/rpt_family="MER1_type"
repeat_region 13488. 13527
/rpt_family="MIR"
repeat_region 15273. 15566
/rpt_family="Alu"
repeat_region 15767. 15829
/rpt_family="MER1_type"
repeat_region 15947. 16238
/rpt_family="Alu"
repeat_region 16244. 16326
/rpt_family="MER1_type?"
repeat_region 16327. 16913
/rpt_family="L1"
repeat_region 16912. 17147
/rpt_family="L1"
repeat_region 17199. 17400
/rpt_family="MAR"
repeat_region 17401. 17517
/rpt_family="MER1_type?"
repeat_region 18465. 18536
/rpt_family="L1"
repeat_region 18644. 18689
/rpt_family="(CA)n"
repeat_region 19101. 19157
/rpt_family="MIR"
repeat_region 19436. 20346
/rpt_family="L1"
repeat_region 20347. 20650
/rpt_family="Alu"
repeat_region 21158. 21226
/rpt_family="L1"
repeat_region 21425. 21738
/rpt_family="Alu"

```


Db 1193 ataataatgaagaagaataattcttatgaaataaagaatgacaaggaggatccca 1252
QY 91 ATGTCGTCATGGTTTAAATATGAAATATGCAATTAAGCGATCTAGACATGCTGTAAA 150
Db 1253 ttcttaagagagataaataatgctgtaaaatgaatgggcttataagaatctgaag 1312
QY 151 AGCAGTGTGGTCCCAATATTAATCTAGCTCTGCAATGGCGGATATCTAATTTGTATA 210
Db 1313 atgatatgataagatataatgataaagaatgatacaaggctggagtaactaaltgata 1372
QY 211 GTTGCTTAAAGAACTCTTATTCGTCGAGTAGAGTTAAAGGCTTTTATCATATCCATNG 270
Db 1373 atataataataaataatgataataatataatgataataatataatataatgata 1432
QY 271 GGTGATTTCTTAACCTTAATAGGGGTTTCGACGAGTATATGGTTGATACATTTCCCTA 330
Db 1433 ataataataataataatgataataatataatataatgataataatataatata 1492
QY 331 CTCACGATTTGTTTCTTTTATCAATTAATAGCATATTTTCATAGCTATATCAATCTTAC 390
Db 1493 atgataataatataatgataataatataatataatataatataatataatata 1552
QY 391 AATCATGATGAGCTGTCTCTAATATAG 421
Db 1553 ataataataataatataatataatctg 1583

RESULT 3
A70227/c
ID A70227 standard; DNA; 4197 BP.
XX AC A70227;
XX DT 07-NOV-2000 (first entry)
DE Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:360.
XX DE
XX DE Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
XX KW antimalarial; malaria; protozoacide; infection; insecticide; ds.
XX OS Plasmodium falciparum.
XX OS
XX PN WO200025728-A2.
XX PD 11-MAY-2000.
XX PE 05-NOV-1999; 99WO-US26796.
XX PR 05-NOV-1998; 98US-0107131.
XX PA (HOFF/) HOFFMAN S.
XX PA (CARU/) CARUCCI D.
XX PA (GARD/) GARDNER M.
XX PA (VENT/) VENTER J C.
XX PI Hoffman S, Carucci D, Gardner M, Venter JC;
XX DR WPI: 2000-365347/31.
XX PT Proteins encoded by chromosome 2 of the human malarial parasite,
XX PT Plasmodium falciparum, useful as antimalarial vaccines and in the
XX PT diagnosis of P.falciparum infection -
XX PS Disclosure; Page 548-549; 577pp; English.
XX CC The present invention describes proteins and their fragments (I) encoded
XX CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
XX CC Also described are: (1) nucleotide sequences (II) encoding (I); and (2)
XX CC vaccines against P. falciparum infection comprising (I) or (II).
XX CC (1) and (II) are useful for the development of vaccines against
XX CC P. falciparum infection. (I) and polyclonal antisera or a monoclonal
XX CC antibody raised to immunogens comprising the sequences of (1), are
XX CC useful in the detection of infection with P. falciparum. Furthermore,

CC (I) (especially when they are rifins or secreted or membrane proteins)
CC can aid the identification of drugs to treat or prevent P. falciparum
CC infection, or they can be used to identify drug resistance in
CC P. falciparum. Sequencing of the Plasmodium chromosome 2 and the
CC subsequent identification of proteins encoded by it will help to expand
CC our understanding of parasite biology, a process hampered by the
CC complexity of the parasite life cycle, and provide new targets for
CC vaccine and drug development. Parasite resistance to drugs and mosquito
CC resistance to insecticides have led to a resurgence of malaria in many
CC parts of the world, and there is a pressing need for vaccines and new
CC drugs. A70078 to A70287 and B18144 to B18352 represent nucleotide and
CC protein sequences given in the present invention, but which are not
CC specifically mentioned within the specification.
XX SQ

Query Match 7.3%; Score 40.6; DB 21; Length 4197;
Best Local Similarity 46.1%; Pred. No. 0.44;
Matches 136; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

QY 125 TAAGGATCTGACATGCTTTAAAGCAGTGTCTCCATATTCATCTAGGCTTG 184
Db 1441 TAACAAATTTATGAATATTCATATAAATAATATATCCAGACGATTTATATTAACATAT 1382
QY 185 CATTGTGGGATACTAATTTTATAGTTGTTTAAAGAAATCTTTCGTGAGTACAG 244
Db 1381 CATTGTGAAATATAAATTTGACAGATGGTTCTATCATATTTTCTTAATATATTA 1322
QY 245 TTAAGCATCTTAAATCATATCCATGCGTGTATTTCTTAACCTTAATAGGGGTTTTCAGC 304
Db 1321 TTTTATGAGCTTTTAAAAAATCGTTAATTAATATTCCTGTTTGTCTATAACAA 1262
QY 305 AAGTATATGTTGATACATTTCTAGTCTCGATTTGTTTCTTATCAATAATAGCAT 364
Db 1261 CATATTTTTCATGATTAAGATTTTCTTTTCTTTTCTTTTCTTATATATATATAT 1202
QY 365 ATTTCATGCTATATATCAATCTTACAAATCATGATGATGCTTCTACTAATAAT 419
Db 1201 TGGATAATTGATAGAGATTTATTTGATGATATCTTATTTCTCTCTTTTAT 1147

RESULT 4
A70101/c
ID A70101 standard; DNA; 3927 BP.
XX AC A70101;
XX DT 07-NOV-2000 (first entry)
DE Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:234.
XX DE
XX DE Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
XX KW antimalarial; malaria; protozoacide; infection; insecticide; ds.
XX OS Plasmodium falciparum.
XX OS
XX PN WO200025728-A2.
XX PD 11-MAY-2000.
XX PE 05-NOV-1999; 99WO-US26796.
XX PR 05-NOV-1998; 98US-0107131.
XX PA (HOFF/) HOFFMAN S.
XX PA (CARU/) CARUCCI D.
XX PA (GARD/) GARDNER M.
XX PA (VENT/) VENTER J C.
XX PI Hoffman S, Carucci D, Gardner M, Venter JC;
XX DR WPI: 2000-365347/31.

Oy		330	AGTCACGATATTGTTTCTTATCATATAATACCAATTTTCANAGCTAATATCATAACTTGA	389
Dd		403	attacaatggttttgataataatagcaaaagcccttggtaagaagtgcataacttga	462
Oy		390	CAAAATCATGATAGCTGTTCCTACTAATATAAGTCATATTTGAA	433
Dd		463	ataacataataaccaccticccttgattgltgtttaacatgcca	506
RESULT	6			
V71729				
ID	V71729	standard; DNA:	5391 BP.	
XX	AC			
XX	V71729;			
XX				
DE	08-FEB-1999	(first entry)		
XX				
DE		Upstream sequence of the ACC synthase gene.		
KW		Apple beta galactosidase; ABG1; ACC; promoter; ripening; fruit; apple;		
KM		1-Aminocyclopropane-1-carboxylate synthase; transgenic; genome mapping;		
RN		transcription factor; ss.		
XX				
OS		Malus sylvestris.		
XX				
PN		MO9845445-A1.		
PD				
PF	15-OCT-1998.			
XX				
EF	03-APR-1998;	98WO-GBO1000.		
XX				
PR	31-MAY-1997;	97GB-0011233.		
XX	09-APR-1997;	97GB-0007193.		
PA		(UKAG-) UK MIN FISHERIES & FOOD.		
XX				
PI	Gillins JR, Hiles ER, James DJ;			
DR	WPI; 1998-583199/49.			
PT		New inducible plant promoters from apple - useful to construct		
PS	vectors for tissue-specific expression of transgenes			
XX				
PS	Claim 10; Fig 5; 52pp; English.			
CC		This represents the upstream region of the 1-Aminocyclopropane-1-		
CC		Carboxylate synthase (ACC) gene and incorporates a promoter sequence.		
CC		The invention provides a new recombinant polynucleotide that comprises a		
CC		promoter sequence which is an inducible promoter obtainable from apple or		
CC		a functional portion of that promoter. The promoter is most preferably		
CC		the Apple beta galactosidase (ABG1) promoter or the ACC synthase promoter		
CC		and is activated in response to tissue specific agents (especially agents		
CC		specific to ripening fruits). A host plant cell transformed with a		
CC		replication vector comprising one of the promoters and a replication		
CC		element which permits replication of the vector can be used in a method		
CC		to produce transgenic apple plants. The transgenic plants can be		
CC		cultivated to produce apples having a modified phenotype. The invention		
CC		may be used to induce protein expression in specific plant tissue, for		
CC		example in order to control fruit ripening, or to introduce protective		
CC		proteins. The invention may also be used to isolate other ripening-		
CC		specific promoters or transcription factors, or in genome mapping. A		
CC		probe comprising one of the promoter sequences can be used to separate a		
CC		transcription factor from fruit cells.		
XX				
SO	Sequence 5391 BP; 1815 A; 929 C; 855 G; 1792 T; 0 other;			
Query Match	7.1%; Score 39.2; DB 19; Length 5391;			
Best Local Similarity	50.0%; Pred. No. 1;			
Matches	98; Conservative 0; Mismatches 98; Indels 0; Gaps 0;			
OY	195 GATTAATATTGGTTATAGTGCTTTAGAATCTTTATTCGTGAGTAGAGTTAAAGATC	254		

Dd		1561 gctcaatttttllttaatgltgtttaacgttaacttaagtlaattaagtlgtttaaatlg	1620
Oy	255 TTTTAAATCAATGCATGCGTGTAATTCTTAACCTTAATAAGGGGTTCGACAAGAATATAGC	314	
	- - - - -		
Dd	1621 tttaatgttgttctaatacgttaactaaagtcacttaaaggtyttcaaagacctagaagcta	1680	
Oy	315 TTGGATCATCTTCCTAGTCTCAGTATATGTTTTCTTATCATATATAGCATATTTCCAAGC	374	
	- - - - - - - - - - - - - - -		
Dd	1681 ttaggaaaaaatatgaattttttaattttttttttaatttgttaaaaataaaataaataat	1740	
Oy	375 TAATCAATACTTGAC	390	
	- - - - -		
Dd	1741 gatgtacaatgtagc	1756	
RESULT	7		
V71738	V71738 standard; DNA; 5407 BP.		
XX AC V71738;			
XX DT 08-FEB-1999	(first entry)		
DE XX Upstream sequence of the ACC synthase gene.			
KW KM Apple beta galactosidase; ABG1; ACC; promoter; ripening; fruit; apple; transcription factor; ss. Malus sylvestris.			
OS MO09845445-AI.			
PN PD 15-OCT-1998.			
XX PF 03-APR-1998;	98MO-GB01000.		
PR PR 31-MAY-1997;	97GB-0011233.		
PP PR 09-APR-1997;	97GB-0007193.		
PA PI (UKAG-) UK MIN FISHERIES & FOOD.			
PI PL Gittins JR, Hiles ER, James DJ;			
DR DR WPI; 1998-58319/49.			
PT PT New inducible plant promoters from apple - useful to construct vectors for tissue-specific expression of transgenes			
PS PS Disclosure: Fig 6; 52pp; English.			
XX CC This represents the upstream region of the l-Aminocyclopropane-l-Carboxylate synthase (ACC) gene and incorporates a promoter sequence. The invention provides a new recombinant polynucleotide that comprises a promoter sequence which is an inducible promoter obtainable from apple or a functional portion of that promoter. The promoter is most preferably the Apple beta Galactosidase (ABG1) Promoter or the ACC Synthase Promoter and is activated in response to tissue specific agents (especially agents specific to ripening fruits). A host plant cell transformed with a replication vector comprising one of the polynucleotides and a replication element which permits replication of the vector can be used in a method to produce transgenic apple plants. The transgenic plants can be cultivated to produce apples having a modified phenotype. The invention may be used to induce protein expression in specific plant tissues, for example in order to control fruit ripening, or to introduce protective proteins. The invention may also be used to isolate other ripening-specific promoters or transcription factors, or in genome mapping. A probe comprising one of the promoter sequences can be used to separate a transcription factor from fruit cells.			
SQ Sequence 5407 BP; 1828 A; 930 C; 854 G; 1795 T; 0 Other;			
XX			

	Query Match	Similarity	7.1%;	Score 39.2;	DB 19;	Length 5407;	
	Best Local Similarity	50.0%;	Pred No. 1:				
	Matches	98;	Conservative	0;	Mismatches	98;	Indels
							Gaps
							0
QY	195	GATACCTAATTGTTTAAAGTAGTGTTTAAGAANAATCTTATTCGTGGAGTAGATTAAAGATC	254				
Dd	1561	gttaatctttttaaagtgcgtttaagtgtaactaaagtcattataattaagtcgtttaacatcg	1620				
QY	255	TTTATTCATATCCATGCGGTATTCTTACTTAATAAGGGTTTTGCAGCAAGTATATGG	314				
Dd	1621	tctaactgctcctcaactgaactaacgtaactaatgcttcaaggacttaggaagta	1680				
OY	315	TTCGATCATATTCCCTAGCTCAGTATTTGTTTTCTATCATGAATAATGACATATTTCATAGC	374				
Dd	1681	taggaaaaaaaaagaacttttaatttttttaatttggtaaaaaataaatal	1740				
QY	375	TATATCAATACTTGCAC	390				
Dd	1741	gatgtcaacaatgtggc	1756				
RESULT	8						
ID	V74335	V74335 standard; DNA; 5423 BP.					
XX	V74335:						
AC	V74335:						
DT	16-MAR-1999	(first entry)					
DE	Staphylococcus aureus contig SEQ ID #24.						
KW	Computer readable medium; vaccine; S.aureus infection; immunodetection;						
KM	cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;						
RN	skin infection; surgical wound infection; scalded skin syndrome;						
XX	Toxic shock syndrome; ds.						
OS	Staphylococcus aureus.						
FH	Key	Location/Qualifiers					
FT	misc.Feature	1021..1080					
FT	/tag= a	"these bases represent a line of missing text in					
FT	/note=	the sequence listing in the specification. They					
FT		are included to maintain the nucleotide numbering					
FT		given in the specification for this DNA sequence"					
misc.Feature	2821..2880						
FT	/tag= b	"these bases represent a line of missing text in					
FT	/note=	the sequence listing in the specification. They					
FT		are included to maintain the nucleotide numbering					
FT		given in the specification for this DNA sequence"					
misc.Feature	4621..4680						
FT	/tag= c	"these bases represent a line of missing text in					
FT	/note=	the sequence listing in the specification. They					
FT		are included to maintain the nucleotide numbering					
FT		given in the specification for this DNA sequence"					
EP786519-A2.							
PN	30-JUL-1997.						
XX	XX						
PF	07-JAN-1997;	97EP-0100117.					
XX	XX						
PR	05-JAN-1996;	96US-0009861.					
XX	XX						
PA	(HUMA-) HUMAN GENOME SCI INC.						
PI	Barash SC,	Choi GH,	Dillon PJ,	Fannon MR,	Kunsch CA;		
PI	Rosen CA;						
XX	WP1: 1997-374922/35.						
DR	XX						

PT Polynucleotide(s) and proteins derived from *Staphylococcus aureus* -
PT stored on computer readable medium and used in the production of
PT anti-*S.aureus* vaccines
XX
PS Claim 1; Page 293-296; 3271pp; English.
XX
XX This sequence represents one of 5191 *Staphylococcus aureus* DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the *S.aureus* DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against *S.aureus* infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC *S.aureus* in a sample. *S.aureus* is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the *S.aureus* DNA sequences contained on the
CC computer readable medium.
XX
XX Sequence 5423 BP: 1563 A; 1053 C; 760 G; 1859 T; 188 other;

Query Match	Similarity	6.9%	Score 38.4	DB 18	Length 5423
Best Local	84	Conservative	52.5%	Pred. No. 1.6	
Matches	84	Conservative	0	Mismatches	76
				Indels	0
				Gaps	0

Query	263	TATCCATGGGTCGATTTCTTAACCTAATAAGGGGTTTCGACGAGATATATGTTGGATAC	322
Db	1294	tattaatgcagaatataccaattactactatgttattgagccgccttaaataccaaac	1353
Qy	323	ATTTCCTAGTCACCATATTTGTTTCTTATCAATATATAGCATATTTGCATAGCTATATCAA	382
Db	1354	agttactatgctcgtgatagcaatcttcaatlaacccaatcttaccgaatacaatltgaa	1413
Qy	383	TACTTGACAAATCATGATGATAGCTGTTCTACTAATAATATAGG	422
Db	1414	tataactgatataaatttttaagtgccttataataacatg	1453

RESULT	9
ID	N60392/c
XX	N60392 standard; DNA; 1612 BP.
XX	
AC	N60392:
XX	
DT	27-JUN-1991 (first entry)
XX	
DE	Sequence encoding the Asparagine-Rich Protein (ARP)
DE	A9319 of Plasmodium falciparum.
XX	
KW	Malaria vaccine; ss.
XX	
OS	Plasmodium falciparum.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
XX	2..1612
XX	/*tag= a
PN	W08606075-A.
XX	
PD	23-OCT-1986.
XX	
PF	11-APR-1986; 86MO-A000092.
XX	
ER	25-JUL-1985; 85AU-0001640.
ER	11-APR-1985; 85AU-0000108.
ER	11-APR-1985; 85AU-0001008.


```

PR 04-AUG-1997; 9705-0054646.
XX
XX (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX
XX Pan Y:
PI
XX WPI: 1999-153692/13.
XX
XX New isolated nucleic acid encoding the new human cytokine Tango-77 -
PT used to inhibit inflammation and to screen for specific modulators
XX
XX Example 5; Figure 4; 226pp; English.
XX
XX X02956-X03048 and X22301-X22304 are overlapping BAC genomic sequences
CC containing alternatively spliced forms of human IL-1ra. Such fragments
CC are used in the method of the invention which describes the isolation of
CC a novel human TANGO-77 encoding nucleic acid and protein. Tango-77 is a
CC member of the cytokine superfamily that is expected to inhibit
CC inflammation by binding to the interleukin-1 receptor (IL-1R). It may
CC also bind to a new receptor so could regulate other cellular processes
CC associated with acute or chronic inflammation, e.g. asthma, chronic
CC myelogenous leukaemia, rheumatoid arthritis, psoriasis and inflammatory
CC bowel disease. It may also induce or suppress interleukins, cytokines and
CC growth factors. Modulators of this protein are used to treat or prevent
CC conditions associated with abnormal levels of inflammation, or activity
CC of IL-1 or its receptor complex.
XX
XX Sequence 28866 BP; 8044 A; 6348 C; 6742 G; 7729 T; 3 other:
SQ
Query Match 6.8%; Score 38; DB 20; Length 28866;
Best Local Similarity 53.3%; Pred. No. 3.1;
Matches 80; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
OY 234 GTGAGTAGAGTTAAAGATCTTTATATATATATATATATATATATATATATAGG 293
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC 3054 GTGCTCTATCTCTACAGCTTTCTTACATACGCTTTATATATATATATATAGCAG 2995
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 294 GGTTCGACGAGATATATATATATATATATATATATATATATATATATATATC 353
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2994 TGTGCTGACAAATATATATATATATATATATATATATATATATATATATATC 2935
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 354 AATATACCATATTTTCAATACATATATATATATATATATATATATATATATAT 383
DB 2934 CTCATATATATATATATATATATATATATATATATATATATATATATATAT 2905
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 12
ID 217016/c
XX 217016 standard; cDNA; 1444 BP.
XX
XX 217016;
XX
XX 12-OCT-1999 (first entry)
XX
DE Human gene expression product cDNA sequence SEQ ID NO:4486.
XX
XX Human; gene; gene expression product; diagnosis; therapy; probe;
KW detection; mapping; tissue typing; profiling; forensic; cancer;
KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX
XX Homo sapiens.
XX
XX MO9938972-A2.
XX
XX 05-AUG-1999.
XX
XX 28-JAN-1999; 99WO-US01619.
XX
XX 03-APR-1998; 98US-0080666.
XX 28-JAN-1998; 98US-0072910.
XX 24-FEB-1998; 98US-0075954.
XX 31-MAR-1998; 98US-0080114.
XX
DE

```

```

PR 03-APR-1998; 9805-0080515.
XX
XX (CHIR ) CHIRON CORP.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Cirvenjakov R, Dickson M, Drmanac R, Drmanac S;
PI Escobedo J, Garcia PD, Garcia V, Glese K, Innis MA;
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
PI Lanson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
XX
XX WPI: 1999-494092/41.
XX
XX Novel human genes and their expression products which are
PT differentially expressed in different cell types
XX
XX Claim 1; Page 2125-2126; 2479pp; English.
XX
XX The present invention describes a library of human polynucleotides
CC comprising the sequences given in 212532 to 217779. Also described is a
CC method of detecting differentially expressed genes correlated with the
CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample from a cell
CC suspected of being cancerous, where the gene product is encoded by one
CC of the 5248 polynucleotide sequences given in 212532 to 217779. The
CC polynucleotides can be used as a source of primers and probes, which can
CC be used for a variety of purpose, e.g. detection of expression levels,
CC mapping, tissue typing or profiling, forensics, genetic analysis and
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
CC can be used for raising antibodies for experimental, diagnostic and
CC therapeutic purposes. The polynucleotides may also be used to construct
CC arrays for diagnostics (which may be used to determine function of an
CC encoded protein); and to detect differences in expression levels between
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC identify a genetic predisposition or susceptibility to a disease such as
CC cancer). The polynucleotides of the invention are especially used in the
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists.
XX
XX Sequence 1444 BP; 534 A; 157 C; 95 G; 209 T; 449 other:
SQ
Query Match 6.8%; Score 37.8; DB 20; Length 1444;
Best Local Similarity 39.5%; Pred. No. 1.7;
Matches 81; Conservative 0; Mismatches 124; Indels 0; Gaps 0;
OY 203 TTGTTATAGTTCGTTAGAAATCTTTATTCGTCGACGAGTTAAAGATCTTTAATCA 262
DB 771 TTATCTTTTGTGAGTAGATATATATATATATATATATATATATATATATATCTN 712
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 263 TATCATGGGTGTATATTTCTTAATATATAGGGGTTTCGACGAGTATATGTTGATAC 322
DB 711 CTNCGATNTGCTNTANATNTATATATATATATATATATATATATATATATATAT 652
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 323 ATTTCCTAGCTCTCACTATGTTTCTTATATCAATATATATATATATATATATATCA 382
DB 651 ANAGCNGNNTNTTATNTCTTTGNNANATATATATATATATATATATATATATGTTNN 592
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 383 TACTTGACAATATCATGATCTGT 407
DB 591 TTNTNGCTNTNGATGNTTGCNNNT 567
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 13
ID A70199/c
XX A70199 standard; DNA; 6042 BP.
XX
XX A70199;
XX
XX 07-NOV-2000 (first entry)
XX
XX Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:332.
XX
DE

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2001, 21:48:22 ; Search time 60.87 Seconds

(without alignments)
1591.991 Million cell updates/sec

Title: US-09-382-242-23

Perfect score: 555
Sequence: 1 AUGCTTTAAACAGACCTC.....AATCATACACTAAGCATAG 555

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 302621 segs, 87301344 residues

Total number of hits satisfying chosen parameters: 605242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgnl_7/ptodata/1/ina/5A.COMB.seq:*\n2: /cgnl_7/ptodata/1/ina/5B.COMB.seq:*\n3: /cgnl_7/ptodata/1/ina/6A.COMB.seq:*\n4: /cgnl_7/ptodata/1/ina/6B.COMB.seq:*\n5: /cgnl_7/ptodata/1/ina/PCUS.COMB.seq:*\n6: /cgnl_7/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	555	100.0	555	2	US-08-602-359A-23
2	38	6.8	152331	4	US-09-128-155-16
3	38	6.8	176373	4	US-09-128-155-17
4	36	6.5	4673	1	US-07-638-431-1
5	36	6.5	4673	5	PCT-US92-00018-1
6	35.6	6.4	1218	2	US-08-731-722-4
7	34.8	6.3	1958	1	US-08-137-175A-7
8	34.8	6.3	1958	4	US-08-479-017-7
9	34	6.1	5775	1	US-08-306-691B-15
10	34	6.1	5775	5	PCT-US93-06251-29
11	33.6	6.1	446	4	US-09-097-541-1
12	33.6	6.1	19124	2	US-08-487-826B-13
13	33.2	6.0	450	1	US-08-090-523-28
14	33.2	6.0	450	1	US-08-398-627-28
15	33.2	6.0	450	1	US-08-406-857-2
16	33.2	6.0	450	1	US-08-596-024-4
17	33.2	6.0	450	5	PCT-US94-07072-2
18	33	5.9	1853	1	US-07-849-438-2
19	32.8	5.9	2014	2	US-08-525-742-7
20	32.8	5.9	2022	2	US-08-505-486-96
21	32.8	5.9	2022	3	US-08-801-028-96
22	32.8	5.9	2022	3	US-09-340-154-96
23	32.8	5.9	2022	5	PCT-US95-09338-96
24	32.8	5.9	2022	5	PCT-US95-09339-96
25	32.8	5.9	2127	2	US-08-505-486-95
26	32.8	5.9	2127	2	US-08-801-028-95
27	32.8	5.9	2127	3	US-09-340-154-95

C 28	32.8	5.9	2127	5	PCT-US95-09338-95	Sequence 95, Appl
C 29	32.8	5.9	2127	5	PCT-US95-09339-95	Sequence 95, Appl
30	32.6	5.9	5910	1	US-08-195-814-1	Sequence 1, Appl
31	32.4	5.8	1186	2	US-08-731-722-5	Sequence 5, Appl
32	32.4	5.8	1617	2	US-08-467-663C-25	Sequence 25, Appl
33	32.4	5.8	1617	2	US-08-838-189D-25	Sequence 25, Appl
34	32.4	5.8	1617	3	US-08-852-344D-25	Sequence 25, Appl
35	32.4	5.8	1617	3	US-08-344-639E-25	Sequence 25, Appl
36	32.4	5.8	1844	2	US-08-467-961A-1	Sequence 25, Appl
37	32.4	5.8	1844	2	US-08-838-189D-1	Sequence 1, Appl
38	32.4	5.8	1844	3	US-08-852-344D-1	Sequence 1, Appl
39	32.4	5.8	1844	3	US-08-344-639E-1	Sequence 1, Appl
40	32.4	5.8	1844	4	US-08-467-969A-1	Sequence 1, Appl
41	32.4	5.8	1844	4	US-08-467-961A-1	Sequence 1, Appl
42	32.2	5.8	5181	1	US-08-257-073-10	Sequence 10, Appl
C 43	32	5.8	1697	1	US-08-472-028A-9	Sequence 9, Appl
C 44	32	5.8	1697	1	US-09-071-286-9	Sequence 9, Appl
45	32	5.8	5117	4	US-08-834-385-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-602-359A-23
Sequence 23, Application US/08602359A
Patent No. 5942430
GENERAL INFORMATION:
APPLICANT: ROBERTSON, Daniel E.
APPLICANT: MORPHY, Dennis
APPLICANT: REID, John
APPLICANT: MAFFIA, Anthony
APPLICANT: LINK, Steven
APPLICANT: SWANSON, Ronald V.
APPLICANT: WARREN, Patrick V.
APPLICANT: KOSMOTKA, Anna
TITLE OF INVENTION: ESTERASES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESS: 4225 EXECUTIVE SQUARE, STE 1400
CITY: LA JOLLA
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602,359A
FILING DATE: February 16, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: HALL, LISA A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 09010/010001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-678-5070
TELEFAX: 619-678-5099
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 555 NUCLEOTIDES
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: GENOMIC DNA
US-08-602-359A-23

Query Match 100.0%; Score 555; DB 2; Length 555;
Best Local Similarity 100.0%; Pred. No. 3.5e-137;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGCTTTAAACAAGCCTCTTGATGATATGATATATTTATCTCAGCTTTCTTC 60
DB 1 ATGCTTTAAACAAGCCTCTTGATGATATGATATATTTATCTCAGCTTTCTTC 60
QY 61 CCATTAAACATGATCCGATAGCATCTATGCGTCATGCTTATATATATGAGATAT 120
DB 61 CCATTAAACATGATCCGATAGCATCTATGCGTCATGCTTATATATATGAGATAT 120
QY 121 GCATTAAAGCATCTAGACATGCTGTTAAAGAGTTCCTCCAAATATTCATCTAGT 180
DB 121 GCATTAAAGCATCTAGACATGCTGTTAAAGAGTTCCTCCAAATATTCATCTAGT 180
QY 181 CTTCGCAATTTGGTGGATGATATGTTATGTTGTTAAGAAATCTTTATCTGAGAT 240
DB 181 CTTCGCAATTTGGTGGATGATATGTTATGTTGTTAAGAAATCTTTATCTGAGAT 240
QY 241 AGAGTTAAAGATCTTAAATCATATCCATGGTGTATTTCTTAACTTAAAGGGTTTC 300
DB 241 AGAGTTAAAGATCTTAAATCATATCCATGGTGTATTTCTTAACTTAAAGGGTTTC 300
QY 301 GACGAGTATATGTTGGATGATACATTTCTAGTCTGATGTTGTTTCTTATCAATATA 360
DB 301 GACGAGTATATGTTGGATGATACATTTCTAGTCTGATGTTGTTTCTTATCAATATA 360
QY 361 GCATTTTCTAGTCTATATCAATATGATGATGATGATGATGATGATGATGATGAT 420
DB 361 GCATTTTCTAGTCTATATCAATATGATGATGATGATGATGATGATGATGATGAT 420
QY 421 GGTCAATATTCGCAATGTTGATGATACATTTGCTTCAAGAGTTCGAGAGGTCGAT 480
DB 421 GGTCAATATTCGCAATGTTGATGATACATTTGCTTCAAGAGTTCGAGAGGTCGAT 480
QY 481 CCCGAGTATTTAGCGTATTTCTGTTTACCATTTCTATATAGACAGTATTTTAAATCA 540
DB 481 CCCGAGTATTTAGCGTATTTCTGTTTACCATTTCTATATAGACAGTATTTTAAATCA 540
QY 541 TACACTAAAGCATAG 555
DB 541 TACACTAAAGCATAG 555
```

RESULT 2

US-09-128-155-16/C
Sequence 16, Application US/09128155
Patent No. 6117654

GENERAL INFORMATION:
APPLICANT: Pan, Yang
TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
FILE REFERENCE: 09404/052001
CURRENT APPLICATION NUMBER: US/09/128,155
EARLIER FILING DATE: 1998-08-03
EARLIER APPLICATION NUMBER: US 60/091,650
EARLIER FILING DATE: 1998-07-02
EARLIER APPLICATION NUMBER: US 60/054,646
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 16
LENGTH: 152331
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(152331)
OTHER INFORMATION: n = A,T,C or G
US-09-128-155-16

Query Match 6.8%; Score 38; DB 4; Length 152331;
Best Local Similarity 53.3%; Pred. No. 0.77;
Matches 80; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

```
QY 234 GTGAGTAGAGTTAAAGATCTTTAATCATATCCATGAGGTCATTTCTTAACTTAATAG 293
DB 137980 GTGAGTAGAGTTAAAGATCTTTAATCATATCCATGAGGTCATTTCTTAACTTAATAG 293
QY 294 GCTTTTGACGAGTATATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 353
DB 137920 TGTGTCGACAAATTTATTTGTAAGATACATTTCTTAACTTATGATGATGATGATGAT 137861
QY 354 AATAATAGCATATTTCTAATAGCTATATCAAT 383
DB 137860 CTCAATATGATATCTCAAAACATGCTTCAT 137831
```

RESULT 3

US-09-128-155-17/C
Sequence 17, Application US/09128155
Patent No. 6117654

GENERAL INFORMATION:
APPLICANT: Pan, Yang
TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
FILE REFERENCE: 09404/052001
CURRENT APPLICATION NUMBER: US/09/128,155
EARLIER FILING DATE: 1998-08-03
EARLIER APPLICATION NUMBER: US 60/091,650
EARLIER FILING DATE: 1998-07-02
EARLIER APPLICATION NUMBER: US 60/054,646
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 17
LENGTH: 176373
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(176373)
OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17

Query Match 6.8%; Score 38; DB 4; Length 176373;

Best Local Similarity 53.3%; Pred. No. 0.8; Mismatches 70; Indels 0; Gaps 0;

```
QY 234 GTGAGTAGAGTTAAAGATCTTTAATCATATCCATGAGGTCATTTCTTAACTTAATAG 293
DB 150561 GTGAGTAGAGTTAAAGATCTTTAATCATATCCATGAGGTCATTTCTTAACTTAATAG 293
QY 294 GCTTTTGACGAGTATATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 353
DB 150501 TGTGTCGACAAATTTATTTGTAAGATACATTTCTTAACTTATGATGATGATGATGAT 150442
QY 354 AATAATAGCATATTTCTAATAGCTATATCAAT 383
DB 150441 CTCAATATGATATCTCAAAACATGCTTCAT 150412
```

RESULT 4

US-07-638-431-1
Sequence 1, Application US/07638431
Patent No. 5198535

GENERAL INFORMATION:
APPLICANT: Hoffman, Stephen L.
APPLICANT: Charoenvit, Yupin
APPLICANT: Hegstrom, Richard
APPLICANT: Khushf, William O.
APPLICANT: Rogers IV, William O.

```

: TITLE OF INVENTION: Protective malaria sporozoite surface protein
: TITLE OF INVENTION: Immunogen and gene
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: A. David Spevack
: STREET: NMDC Building 1 T-12 National Naval
: CITY: Bethesda
: STATE: MD
: COUNTRY: USA
: ZIP: 20814-5044
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.24
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/638,431
: FILING DATE: 19910110
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Spevack, Avram D.
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 295-6759
: TELEFAX: (301) 295-4033
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4673 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: N
: ANTI-SENSE: N
: ORIGINAL SOURCE:
: ORGANISM: Plasmodium yoelii
: STRAIN: 17X(NL)
: DEVELOPMENTAL STAGE: erythrocytic stage
: TISSUE TYPE: Blood
: CELL TYPE: erythrocytic stage
: IMMEDIATE SOURCE:
: LIBRARY: Py-lambdaagtl1-2-7 kb genomic expression
: CLONE: Py10.1111
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 718..3195
: OTHER INFORMATION:
:
: US-07-638-431-1
:
: Query Match 6.5%; Score 36; DB 1; Length 4673;
: Best Local Similarity 64.3%; Pred. No. 0.93;
: Matches 54; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
:
: QY 296 TTTTCGACGAAGTATATGCTGATACATTTCTAGCTCAGTATTGTTTCTTATCAA 355
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: DB 263 TTTTCGACGAAGTATATGCTGATACATTTCTAGCTCAGTATTGTTTCTTATCAA 322
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: QY 356 TAATGCATATTTCTAGCTATAT 379
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: DB 323 TATTAACATTCATAAAATATATAT 346
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
:
: RESULT 5
: PCT-US92-00018-1
: Sequence 1, Application PC/TUS9200018
: GENERAL INFORMATION:
: APPLICANT: Hoffman, Stephen L.
: APPLICANT: Charoenvit, Yupin
: APPLICANT: Hedstrom, Richard
: APPLICANT: Khumritth, Srisin
: APPLICANT: Rogers IV, William O.
: TITLE OF INVENTION: Protective malaria sporozoite surface protein

```

```

: TITLE OF INVENTION: Immunogen and gene encoding
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: A. David Spevack
: STREET: NMDC Building 1 T-12 National Naval
: CITY: Bethesda
: STATE: MD
: COUNTRY: USA
: ZIP: 20814-5044
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.24
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US92/00018
: FILING DATE: 19920103
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Spevack, Avram D.
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 295-6759
: TELEFAX: (301) 295-4033
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4673 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: N
: ANTI-SENSE: N
: ORIGINAL SOURCE:
: ORGANISM: Plasmodium yoelii
: STRAIN: 17X(NL)
: DEVELOPMENTAL STAGE: erythrocytic stage
: TISSUE TYPE: Blood
: CELL TYPE: erythrocytic stage
: IMMEDIATE SOURCE:
: LIBRARY: Py-lambdaagtl1-2-7 kb genomic expression
: CLONE: Py10.1111
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 718..3195
: OTHER INFORMATION:
:
: PCT-US92-00018-1
:
: Query Match 6.5%; Score 36; DB 5; Length 4673;
: Best Local Similarity 64.3%; Pred. No. 0.93;
: Matches 54; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
:
: QY 296 TTTTCGACGAAGTATATGCTGATACATTTCTAGCTCAGTATTGTTTCTTATCAA 355
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: DB 263 TTTTCGACGAAGTATATGCTGATACATTTCTAGCTCAGTATTGTTTCTTATCAA 322
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: QY 356 TAATGCATATTTCTAGCTATAT 379
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: DB 323 TATTAACATTCATAAAATATATAT 346
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
:
: RESULT 6
: US-08-731-722-4
: Sequence 4, Application US/08731722
: Patent No. 5961971
: GENERAL INFORMATION:
: APPLICANT: Martin, Frank N.
: TITLE OF INVENTION: Blocontrol of Fungal Solborne Pathogens
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Saliwanhik & Saliwanhik

```

STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606-6669
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/731.722
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF-161
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: 17-1
US-08-731-722-4

Query Match 6.4%; Score 35.6; DB 2; Length 1218;
Best Local Similarity 48.1%; Pred. No. 0.8;
Matches 101; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 150 AAGCAGTGTGCTCCAAATTCATAGCTTGCATTTGGGATGACTAATGTTAT 209
DB 512 AACTTGTGTAATCCACCAAAATTTAACTATTAATGAAGACTTTTACTTAT 571

QY 210 AGTTGGTTAAGAAATTTTATCGTAGAGTAAGACTTAAAGACTTAAATCAT 269
DB 572 ATTCTTAATTAATATCTGAATTTAAACAGAGCCTTAAGACCTTAATTAAGCTAT 631

QY 270 GGGTATTTCTTAACCTAATAGGGCTTTTCGACGAGTATATGTTGATACATTTCT 329
DB 632 AAGAAATTTTTCATAGAAAGSTTTTATTTTAAATGAAATTTGTAATCTTATTT 691

QY 330 AGTCTCAGTATGTTTCTTATCATTAAT 359
DB 692 TACTTTAATAAGAAATTAATTAATTAAT 721

RESULT 7
US-08-137-175A-7
Sequence 7, Application US/08137175A
Patent No. 5777095
GENERAL INFORMATION:
APPLICANT: BARBOUR, Alan G.
APPLICANT: BERGSTROM, Sven
APPLICANT: HANSSON, Lennart
TITLE OF INVENTION: IMPROVEMENT IN BORRELIA BURGDOFFERI AND
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/137.175A
FILING DATE: 26-Oct-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/08972
FILING DATE: 22-Oct-1992
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, Iver P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: BARBOUR-1B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Borrelia burgdorferi
STRAIN: 1p90
INDIVIDUAL ISOLATE: Isolate from I. persulcatus from
INDIVIDUAL ISOLATE: Soviet Union
FEATURE:
NAME/KEY: CDS
LOCATION: 125..949
OTHER INFORMATION: /product= "Ospa"
NAME/KEY: CDS
LOCATION: 959..1843
OTHER INFORMATION: /product= "Ospb"
US-08-137-175A-7

Query Match 6.3%; Score 34.8; DB 1; Length 1958;
Best Local Similarity 48.5%; Pred. No. 1.5;
Matches 96; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 196 ATACTAATTCCTATAGTTCCTTAAGAAATCTTATTCGAGTAGAGTAAGATCT 255
DB 5 ATATTTGAATTAATTAATCATTTTATTTTATTTTATTTGCTATTTGTTGATCT 64

QY 256 TTAATCATATCCATGGGTATTTCTTAACCTAATAGGGGTTTCGACGAGTATATGT 315
DB 65 TATACTAATTAATTAATTTGTTATTAAGTTATTAATTAATTAATTAAGGAAATTAAT 124

QY 316 TGGATACATTTCTAGTCTGATGTTTCTTATTAATTAAGCATATTTATATGCT 375
DB 125 ATGAAAAAATTAATTTAGGCAATAGCTAATTAATTAATTAATTAATTAATTAAT 184

QY 376 ATATCAATACCTTGACAAA 393
DB 185 GTTAGACGCTTGATGAA 202

RESULT 8
US-08-479-017-7
Sequence 7, Application US/08479017
Patent No. 6143872
GENERAL INFORMATION:
APPLICANT: BARBOUR, Alan G.
APPLICANT: BERGSTROM, Sven
APPLICANT: HANSSON, Lennart
TITLE OF INVENTION: IMPROVEMENT IN BORRELIA BURGDOFFERI AND
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:


```
RESULT 10
PCT-US93-06251-29
; Sequence 29, Application PC/TUS9306251
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric and Rife, Jason P.
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
;   STEREOISOMERICALLY PURE ALKYLPHOSPHONATES AND ARYLPHOSPHONATES
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; City: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06251
; FILING DATE: 19930630
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digilio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8586
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5775 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PCT-US93-06251-29

Query Match          6.1%; Score 34; DB 5; Length 5775;
Best Local Similarity 44.8%; Pred. No. 3.3; Mismatches 160; Indels 0; Gaps 0;
Matches 130; Conservative 0;

QY 4 TCTTTAAACAGCAGCTCTGGATGATATGATATTTATTCACGCTTTTCCTTCCCA 63
DB 4538 TTTGTCAAAAAGTAATATTTCTTGATATTTGTAGTATGTTTGTAGAACCCAGCA 4597
QY 64 TTAACAATGATCCGATAGCTATCTCTATGTCGATGCTTAAATATATGGAATATGCA 123
DB 4598 GTTAACCTTGAAGCTGAAATTTATTTAGTAACCTTGTGTATTAAGTATGATGAA 4657
QY 124 TTAAGCATCTAGACATGCTGTTAAAGCAGTCTGCTCCAAATATCAATCTAGCTT 183
DB 4658 TTTCTGCAATGAGAAAGCAATAGCTGTCATTAATGCTTTCTTCCAAAGAAAGACT 4717
QY 184 GCAATTTGGTGGATCTAATGTTATGTTAGTTAGTAAAGAAATCTTTTCGCTGAGTGA 243
DB 4718 CACATGAGTCTTGAAGAATATGATATAGTAAAGTATGATGCTGTTTATGTTAATAGT 4777
QY 244 GTTAAAGGATCTTAAATCAATATCCATGCGGTGATTTTCTTAACTTAATAGC 293
DB 4778 TTTGAAGTCCCTGTTGGATATGATAGTAATTTAGATCAATTTAGGCG 4827

RESULT 11
US-09-097-541-1
; Sequence 1, Application US/09097541
; Patent No. 6100092
; GENERAL INFORMATION:
; APPLICANT: Borisyuk, Mykola
```

```
APPLICANT: Borisyuk, Lyudmyla
; APPLICANT: Raskin, Ilya
; TITLE OF INVENTION: Materials and methods for amplifying
;   TITLE OF INVENTION: polynucleotides in plants
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; City: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/097,541
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Merkel, William K.
; REGISTRATION NUMBER: 40,725
; REFERENCE/DOCKET NUMBER: 29155/34699
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 446 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "aps rdna sequence"
; US-09-097-541-1

Query Match          6.1%; Score 33.6; DB 4; Length 446;
Best Local Similarity 46.0%; Pred. No. 2; Mismatches 134; Indels 0; Gaps 0;
Matches 114; Conservative 0;

QY 196 ATACTAATGTTAAGTGGTTTAAAGAAATCTTTATTCGAGTAGAGTTAAAGATCT 255
DB 126 ATTTATGATGATGTTGATTTTAAATAGTTTATCGTACTGTTTAAATATTTT 185
QY 256 TTAATCATATCCATGGGTTATTTCTTAATATAGGGGTTTCGACGAAGTATAGGT 315
DB 186 ATTTATGATGTTAATATTAATTAATGATGATGATGAAATTTCTCCATGTTTCT 245
QY 316 TGATACATTTCCAGTCTCAGATTTGTTTCTTATCAATATAGCATATTTCAAGCT 375
DB 246 ATATTATTAATATTTCTTTATTTTGTATTTATATATGATATTTTGTGTTTAAAT 305
QY 376 ATATCAATCTTGACAAATCATGATGATGCTGCTTACTAATATAGTCATATTCGATG 435
DB 306 AATATTTATTAATAAATAATATTTTGTAAATAATATATCATTTACAAATGTTTAAAG 365
QY 436 TGTATCT 443
DB 366 TCATTTGT 373

RESULT 12
US-08-487-826B-13/C
; Sequence 13, Application US/08487826B
; Patent No. 593827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chlenis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
```



```
APPLICANT: Stark, David M.
TITLE OF INVENTION: Enhanced Starch Biosynthesis
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESS: Grace L. Bonner, Monsanto Co. BBAF
STREET: 700 Chesterfield Parkway No. 5608149th
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/398,627
FILING DATE: 03-MAR-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/090,523
FILING DATE: 12-JUL-1993
APPLICATION NUMBER: US 07/709663
FILING DATE: 07-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/539763
FILING DATE: 18-JUN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Bonner, Grace L.
REGISTRATION NUMBER: 32,963
REFERENCE/DOCKET NUMBER: 38-21(10559)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 537-7286
TELEFAX: (314) 537-6047
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-398-627-28

Query Match
Best Local Similarity 50.0%; Score 33.2; DB 1; Length 450;
Matches 83; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 277 TTCTTAACTTAATAGGGCTTTTCGACGAAATATATGTTGATACATTTCTAGTCTCA 336
DB 48 TTCTTTTAAATGATTTCTTCTACTTAATTTGAATTTGGAGCAACTTAATGTAATA 107
QY 337 GTATTGTTTCTTATCAATATAGCATATTTTCATAGCTATATCAATCTTGACAAATCA 396
DB 108 TTTTCTTTTCTTATCAAAATGATGCTGCTATATATAATATCAATGTTATATACA 167
QY 397 TGCATAGCTGTTCTACTAATAATAGTCATATTTGCAATGTGCTATC 442
DB 168 TAAATTTAATATTTTATATAAAAAAATATCGAGCTAATCATATC 213

RESULT 15
US-08-406-857-2
Sequence 2, Application US/08406857
Patent No. 5608150
GENERAL INFORMATION:
APPLICANT: Conner, Timothy W.
TITLE OF INVENTION: Fruit Specific Promoters
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Grace L. Bonner, Monsanto Company, BBAF
STREET: 700 Chesterfield Parkway No. 5608150th
CITY: St. Louis
```

```
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/406,857
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/07072
FILING DATE: 27-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/090,523
FILING DATE: 12-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bonner, Grace L.
REGISTRATION NUMBER: 32,963
REFERENCE/DOCKET NUMBER: 38-21(10655)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-7286
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-406-857-2
```

```
Query Match
Best Local Similarity 50.0%; Score 33.2; DB 1; Length 450;
Matches 83; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 277 TTCTTAACTTAATAGGGCTTTTCGACGAAATATATGTTGATACATTTCTAGTCTCA 336
DB 48 TTCTTTTAAATGATTTCTTCTACTTAATTTGAATTTGGAGCAACTTAATGTAATA 107
QY 337 GTATTGTTTCTTATCAATATAGCATATTTTCATAGCTATATCAATCTTGACAAATCA 396
DB 108 TTTTCTTTTCTTATCAAAATGATGCTGCTATATATAATATCAATGTTATATACA 167
QY 397 TGCATAGCTGTTCTACTAATAATAGTCATATTTGCAATGTGCTATC 442
DB 168 TAAATTTAATATTTTATATAAAAAAATATCGAGCTAATCATATC 213
```

Search completed: May 30, 2001, 22:40:51
Job time: 3149 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2001, 21:07:06 ; Search time 1108.56 Seconds
(without alignments)
4373.723 Million cell updates/sec

Title: US-09-382-242-23

Perfect score: 555

Sequence: 1 ATGCTTTTAAACAAGCACTC.....AATCATACACTTAAACGATAG 555

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9623517 seqs, 4368049070 residues

Total number of hits satisfying chosen parameters: 19247034

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: gb_estba:*
34: gb_estfun:*
35: gb_esthum1:*
36: gb_esthum2:*
37: gb_esthum3:*
38: gb_esthum4:*
39: gb_esthum5:*
40: gb_esthum6:*
41: gb_esthum7:*
42: gb_esthum8:*
43: gb_esthum9:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estp11:*
73: em_estp12:*
74: em_estp13:*
75: em_estp14:*
76: em_estp15:*
77: em_estp16:*
78: em_estp17:*
79: em_estp18:*
80: em_estp19:*
81: em_estp110:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

117: gb_est48:*
 118: gb_est49:*
 119: gb_est50:*
 120: gb_est51:*
 121: gb_est52:*
 122: gb_est53:*
 123: gb_est54:*
 124: gb_est55:*
 125: gb_est56:*
 126: gb_est57:*
 127: gb_est58:*
 128: gb_est59:*
 129: gb_est60:*
 130: gb_est61:*
 131: gb_est62:*
 132: gb_est63:*
 133: gb_est64:*
 134: gb_est65:*
 135: gb_est66:*
 136: gb_est67:*
 137: gb_est68:*
 138: gb_est69:*
 139: gb_est70:*
 140: gb_est71:*
 141: gb_est72:*
 142: gb_est73:*
 143: gb_est74:*
 144: gb_est75:*
 145: gb_est76:*
 146: gb_est77:*
 147: gb_est78:*
 148: gb_est79:*
 149: gb_est80:*
 150: gb_est81:*
 151: gb_est82:*
 152: gb_est83:*
 153: gb_est84:*
 154: gb_est85:*
 155: gb_est86:*
 156: gb_est87:*
 157: gb_est88:*
 158: gb_est89:*
 159: gb_est90:*
 160: gb_est91:*
 161: gb_est92:*
 162: gb_est93:*
 163: gb_est94:*
 164: gb_est95:*
 165: gb_est96:*
 166: gb_est97:*
 167: gb_est98:*
 168: gb_est99:*
 169: gb_est100:*
 170: gb_est101:*
 171: gb_est102:*
 172: gb_est103:*
 173: gb_est104:*
 174: gb_est105:*
 175: gb_est106:*
 176: gb_est107:*
 177: gb_est108:*
 178: gb_est109:*
 179: gb_est110:*
 180: gb_est111:*
 181: gb_est112:*
 182: gb_est113:*
 183: gb_est114:*
 184: gb_est115:*
 185: gb_est116:*
 186: gb_est117:*
 187: gb_est118:*
 188: gb_est119:*
 189: gb_est120:*

Result No.	Score	Query Match	Length	ID	Description
190:	49.6	8.9	1101	CNS00EVL	AL069706 Drosophila
191:	47.4	8.5	1101	CNS00YWL	AL096927 Drosophila
192:	46.6	8.4	1101	CNS000D1	AL065414 Drosophila
193:	46.6	8.4	1101	CNS0039G	AL063921 Drosophila
194:	46.2	8.3	878	CNS0187R	AL108993 Drosophila
195:	45.4	8.2	500	AU086529	AU086529 Drosophila
196:	45.4	8.2	792	BF616219	BF616219 HVMSC000
197:	45.4	8.2	1101	CNS0106X	AL098595 Drosophila
198:	45.2	8.1	1101	CNS0161L	AL106896 Drosophila
199:	45	8.1	641	AQ946120	AQ946120 Sheared D
200:	44.8	8.1	982	CNS013ML	AL103027 Drosophila
201:	44.8	8.1	988	CNS0072R	AL066743 Drosophila
202:	44.6	8.0	491	FR0042156	AL129648 Fugu rubr
203:	44.6	8.0	502	FR0042075	AL129567 Fugu rubr
204:	44.4	8.0	500	AU086530	AU086530 Drosophila
205:	44.4	8.0	856	AZ687828	AZ687828 ENTLE61TR
206:	44.4	8.0	911	AZ677405	AZ677405 ENTJN1TR
207:	44.4	8.0	1146	CNS021G2	AL176843 Tetradon

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

[illegible]

QY 40

DQ 463 CCGAGAGTGTCTATTCCGACTTAATAGCGGTATTCGTTTTACCTCATATA 522
| : | : | : ||| : ||||:
603 AATTAAATTTATATATAATATATATATTTWTAAATTTAATTATTTATTTTAAAATATMTT 544

523 AGACAGTATTTTAAATCATACACTAAACGATA 554

```

Db      543  TTTT TTTT AWTAAWWTWAAAWAATWAMWMTWNA 512
          ||::|::| | | |::|::|

```

RESULT 2

LOCUS	DEFINITION
CNS000YWL	1101 bp DNA GSS 26-JUL-1999
	Drosophila melanogaster genome survey sequence SP6 end of BAC
	BACN01G13 of DrosBAC library from Drosophila melanogaster (fruit

ACCESSION	AL096927
VERSION	AL096927.1
	GI:5608538

ORGANISM

REFERENCE AUTHORS TITLE JOURNAL	COMMENT
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 1101) Genoscope.	
Direct Submission Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequefegenoscope.cns.fr - Web : www.genoscope.cns.fr)	
Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billard at CPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC	

```

FEATURES
source      Location/Qualifiers
1..1101     /organism="Drosophila melanogaster"

```

/plasmid="pBelOBAC11"
/db_xref="taxon:7227"
/clone_11b="DrosBAC"
/clone="BACN01G13"
/note="end : SP6"
BASE COUNT 424 a 119 c 129 g 248 t 181 others
ORIGIN

Query Match 8.5%; Score 47.4; DB 229; Length 1101;
Best Local Similarity 35.1%; Pred. No. 0.11;
Matches 98; Conservative 46; Mismatches 135; Indels 0; Gaps 0;

OY 106 AATATATGGAATATGATTAAGCGATGAGACATGCTGTTAAAGCAGTGGCTCA 165
DB 1100 HAAAT 1041
OY 166 AATATCAATCTAGCTTCCTCAATGCTGAGATACATATGTTATAGTGGTTAACAAT 225
DB 1040 AATTAAT 981
OY 226 CTATATCTGAGAGATAGATTAAGATCTTAATCATATCATGAGTGTATTTCTTAC 285
DB 980 AAT 921
OY 286 TTAATAGGGGTTTCGACAGATATATATGTTGATATCTTCTACTGCTCAATGTT 345
DB 920 TTAAT 861
OY 346 TTCTATCAT 384
DB 860 TTTTAT 822

RESULT 3

CNS000D1

LOCUS 1101 bp DNA GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #

BACR01J16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
AL065414.1 GI:4938827

ACCESSION

VERSION 1
KEYWORDS GSS.

SOURCE

ORGANISM

Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazuoto Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

1. .1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"

/clone_11b="RPCI-98"
/clone="BACR01J16"
/note="end : TET3"
BASE COUNT 280 a 104 c 123 g 211 t 383 others
ORIGIN

Query Match 8.4%; Score 46.6; DB 229; Length 1101;
Best Local Similarity 14.0%; Pred. No. 0.17;
Matches 63; Conservative 203; Mismatches 185; Indels 0; Gaps 0;

OY 21 TTGGATGATATGATATATATATATATATATATATATATATATATATATAT 80
DB 647 KKKAKRAATDDTATATATATATATATATATATATATATATATATATATAT 706
OY 81 AACTATCTATGCTGCTGCTTAAATATATATATATATATATATATATATATAT 140
DB 707 KKKAKRAATDDTATATATATATATATATATATATATATATATATATATAT 766
OY 141 TCGTGTAAAGCAGTGTGCTCAATATATATATATATATATATATATATATAT 200
DB 767 GPKRRRTAGRGDKTKGKKTKGDKGSDGSDGSDGSDGSDGSDGSDGSDGSD 826
OY 201 AATGTTATATAGTGTGTTAAGAAATCTTATTCGTGAGCTAGACGTTAAAGATCTTTAAT 260
DB 827 TGGAKDKAKKRAKRRKRWTRKTRDPAWKAAWAAWRRRRRKRDRRRRRR 886
OY 261 CATATCATGAGGATATATCTTAACTATATAGGGGTTTCGAGCAAGTATATGTTGAT 320
DB 887 GADKKKKKDDDDTAKDDTKKTDITAGTACGDDDKAKGDKKKTDRWKRRKRG 946
OY 321 ACATTTCTAGTCTAGTATGTTTCTTATATATATATATATATATATATATATAT 380
DB 947 RKGDKRKTAGRAGDGGDKKKDKDRTDDAKATGDKDITDPTAKRRRRRRRRGR 1006
OY 381 AATCTTGACAAATCATGATAGCTGTTCTACTAATATATAGCTATATGCAATGTGTA 440
DB 1007 GDRKDKKKRRARRRRKRRKRRKTRDDDDDKRNGRAKAKADAKADADAADA 1066
OY 441 TCTACACTTGGCTTCAGATTCGAGAGT 471
DB 1067 KAKAKDKAKKAKADADDKDKKAKAKRDK 1097

RESULT 4

CNS0039G/c

LOCUS 1101 bp DNA GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #

BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
AL063921
AL063921.1 GI:4941778

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazuoto Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial

REFERENCE	Muscomorphia; Ephydroidea; Drosophilidae;Drosophila.
AUTHORS	1 (bases 1 to 1101)
TITLE	Drosophila
JOURNAL	Genoscope Direct Submission Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr Web : www.genoscope.cns.fr) determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk . This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billard at CEPH (Centre d'Etude du Polymorphism Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelosBAC11. Location/Qualifiers 1..1101 /organism="Drosophila melanogaster" /plasmid="pBelosBAC11" /db_xref="taxon:7227" /clone_lib="DrosBAC" /clone="BACN03K20" /note="end : T7"
COMMENT	
FEATURES	
source	
BASE COUNT	258 a 107 c 60 g 175 t 501 others
ORIGIN	
Query Match	8.2%; Score 45.4; DB 229; Length 1101;
Best Local Similarity	17.8%; Pred. No. 0.34;
Matches	80; Conservative 198; Mismatches 165; Indels 6; Gaps 2;
OY	25 ATGATAGTAAATAATTATTTCTCGACTTTTCCTCCATTACAATGCATTAGCT 84
Db	650 AATATKAARAKARKDKADAMWMDKKAKAKAKAKAKAKAKAKAKATGTADKARKAK 709
OY	85 ATCCTATGCTGCTGGT--TAATATATGAATAATGCAATTAAGCATCTGAGCATGC 143
Db	710 AKAKATTKAKARAWDTAWTATPAADAADKDGKADAKAKAKADADKRDKRWDKJKRRA 769
OY	144 TGTTAAAAGCAGTGTTGCCAATTTCAAATCTAGCTTTGCAATTTGGGAGTACTAAT 203
Db	770 KAKAKAADDAADAKKADAKAKADADDGRGGKKRRRADRKKKKKKKDDAKMGDK 829
OY	204 TGTT----ATAGTTGTTTAAAGAATCTTTATTTCTGSGAGTAGATTAAAGATCTTTA 258
Db	830 KKAKKDAAAAKADAADAKARRRADDKPKADAKAKAKAKAKDDDDAKAATKAKKA 889
OY	259 ATCATATCCATGGGGTATTTCTTAATTAAGGGTTTTGACGAAGTATATGTTGG 318
Db	890 TKAKDKAKAKKKKKKKKDKADAKAKKADADKDDKDDKDKAKDKKKWKDKRAKD 949
OY	319 ATAATTTCTAGTCTGATTTGTTTTCTTATCAATAATGCAATTTTCATTACTATA 378
Db	950 KAKKDDKKDKAKDKDRKKDADAKAKAKADAKAKAKAKAKAADDADAADAKAKAA 1009
OY	379 TCATATCTTGCACAAATCATGATAGCTGTTCTTAATTAATAGTCATATTTGCAATGTGG 438
Db	1010 DAKAKAKAKADDAKAKAKADADADAKAKAKARAARAKAKADADAANAKAKAADKDXDA 1069
OY	439 TAATCACACTTTGCTTCGAGATTCGAG 467
Db	1070 KKDDAKKAKAKAKADAKAKAKAKAKAK 1098
RESULT	9
CNSO16L1	
LOCUS	CNSO16L1 1101 bp DNA GSS 26-Jul-1999
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC
ACCESION	BACNID22 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
VERSION	AL106896
KEYWORDS	GSS.
	GI:5624374

SOURCE	fruit fly.
ORGANISM	Plasmodium falciparum
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
AUTHORS	Piercygott, Neoptera; Endopterygota; Diptera; Brachycera;
JOURNAL	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
TITLE	1 (bases 1 to 1101)
COMMENT	Genoscope. Direct Submission Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr) Web : www.genoscope.cns.fr Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billard at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBeloBAC11.
FEATURES	Location/Qualifiers
SOURCE	1..1101 /organism="Drosophila melanogaster" /plasmid="pBeloBAC11" /db_xref="taxon:7227" /clone_lib="DrosBAC" /clone="BAC11D22" /note="end : 17"
BASE COUNT	203 a 220 c 84 g 158 t 436 others
ORIGIN	
Query Match	8.1% Score 45.2; DB 229; Length 1101; Best Local Similarity 30.6%; Pred. No. 0.38; Matches 90; Conservative 68; Mismatches 136; Indels 0; Gaps 0;
OY	167 TATTCATCTGAGTGGTTCCAACTTGCGGAGACTAATGTATTAAGTGGTTAACAATC 226
DB	775 TTATATTGTTTAKDTTTTTTTTWTATATTTTWTGTTTATTTTMTATATATTTTW 834
OY	227 TTTATTCCTGAGTAGCTTAAAGAGCTTATCAATCATGCCAGTCTATCTTAACT 286
DB	835 AWTATAATTAATATTTTATTAATATATATATATATATATATATATATATATAT 894
OY	287 TAATAGGGGTTTGCAGCAAGTATATGGTGGATACATTTCCCTAGCTCAGTATGTTT 346
DB	895 WATWATATTATWTAAMWTAAMWMAATTTATWTTATATDMRTTASTATATTTTDHW 954
OY	347 TCCTATCAATATAGCATATTTCAAGCTATATCAATCAATCAATCAATCAATCAATC 406
DB	955 TWATATRAKAAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAW 1014
OY	407 TTCCTACTATATAGTGCATATTTGCAATGCTGATCTACACTTTGCTTCAGAGA 460
DB	1015 RMAWADATTTAMDAAATRTAATATWTTTKWTTTKWTTTKWTTTKAKAKRGR 1068
RESULT	10
LOCUS	AQ946120 641 bp DNA GSS 27-JAN-2000
DEFINITION	Sheared DNA -46J23.TR Sheared DNA Trypanosoma brucei genomic clone
VERSION	AQ946120
KEYWORDS	AQ946120.1 GI:6769385
SOURCE	GSS.
ORGANISM	Trypanosoma brucei. Trypanosoma brucei Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
REFERENCE	1 (bases 1 to 641) El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Sub,E., Malek,J., Fujii,C., Gerard,C., Leech,V., de Jong,P., Ullu,E., Melville,S., Donelson,J., Fraser,C. and Adams,M. Determination of clone end sequences from Trypanosoma brucei GUTat
TITLE	

THIS PAGE BLANK (USPTO)